

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:03:10 ; Search time 117.463 Seconds
(without alignments)
184.386 Million cell updates/sec

Title: us-09-936-956-1
Perfect score: 98
Sequence: 1 XGXXXXXXXXXXXXXAX.....XXXXXXXXXXXXXGXXXXXG 56

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep16Dec04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	25.5	260	7 ADE44814	Ade44814 Neisseria
2	25	25.5	260	7 ADE44818	Ade44818 Neisseria
3	25	25.5	261	7 ADE44816	Ade44816 Neisseria
4	25	25.5	274	8 ADP70329	Adp70329 Neisseria
5	25	25.5	1026	2 AAR48993	Aar48993 rsaA S-ly
6	25	25.5	1026	2 AAR94014	Aar94014 Caulobact
7	25	25.5	1026	2 AAW37490	Aaw37490 Caulobact
8	25	25.5	1026	3 AAY44757	Aay44757 Caulobact
9	25	25.5	1026	6 AAE34374	Aae34374 Caulobact
10	23	23.5	8	ADP80011	Adp80011 Human HLA
11	23	23.5	9	ADD23187	Add23187 Breast ca
12	23	23.5	9	ADM72935	Adm72935 Human CEA
13	23	23.5	10	2 AAY46550	Aay46550 Immunogen
14	23	23.5	10	2 AAY00829	Aay00829 HLA-A24 a
15	23	23.5	10	2 AAY00830	Aay00830 HLA-A24 a
16	23	23.5	10	2 AAY00833	Aay00833 HLA-A24 a
17	23	23.5	10	2 AAY00835	Aay00835 HLA-A24 a
18	23	23.5	10	2 AAY00809	Aay00809 HLA-A24 a
19	23	23.5	10	2 AAY00831	Aay00831 HLA-A24 a
20	23	23.5	10	2 AAY00828	Aay00828 HLA-A24 a
21	23	23.5	10	2 AAY00834	Aay00834 HLA-A24 a
22	23	23.5	10	2 AAY00832	Aay00832 HLA-A24 a
23	23	23.5	10	5 AAU82066	Aau82066 CEA antig
24	23	23.5	10	7 ADF83690	Adf83690 HLA-A24 r
25	23	23.5	10	7 ADG38740	Adg38740 Human car

ALIGNMENTS

RESULT 1
ADE44814
ID ADE44814 standard; protein; 260 AA.
XX ADE44814;
AC AC
DT 29-JAN-2004 (first entry)
XX
DE Neisseria meningitidis ORF2086 protein SeqID248.
XX
KW ORF2086; Neisseria meningitidis serogroup B infection; antibacterial;
KW antiinflammatory; immune response; bacterial meningitis;
KW Streptococcus pneumoniae infection; non-pathogenic;
KW immunogenic composition; 2086 protein.
XX
OS Neisseria meningitidis.
PN WO2003063766-A2.
XX
PD 07-AUG-2003.
XX
PF 11-OCT-2002; 2002WO-US032369.
XX
PR 11-OCT-2001; 2001US-0328101P.
PR 30-AUG-2002; 2002US-0406934P.
XX
(AMHP) WYETH HOLDINGS CORP.
PI Zlotnick GW, Fletcher LD, Farley J, Bernfield LA, Zagursky RJ;
PI Metcalf BJ;
XX
DR WPI; 2003-663416/62.
DR N-PSDB; ADE44813.
XX
PT Composition comprising crossreactive immunogenic antigen encoded by open
PT reading frame 2086 of Neisseria sp.; that provides immunogenicity against
PT meningitis, or its immunogenic portion or biological equivalent.
PS
PS Claim 60; SEQ ID NO 248; 480pp; English.
XX
CC This invention relates to a novel composition which comprises at least
CC one protein (or fragment of) encoded by an open reading frame (ORF) of a
CC Neisseria sp. (ORF2086), where the ORF encoding a crossreactive
CC immunogenic antigen provides immunogenicity against infection by
CC Neisseria meningitidis serogroup B in a subject. The composition of the
CC invention may have antibacterial or antiinflammatory activity through the
CC induction of the immune response. The invention may be useful for the
CC treatment of bacterial meningitis in a mammal. One or more polypeptides

26 23 23.5 10 7 ADG38664 Human car
27 23 23.5 10 8 AAO24401 HLA-A24-r
28 23 23.5 10 8 ADMI2345 MHC class
29 23 23.5 10 8 ADO38565 Carcinom
30 23 23.5 10 8 ADP80010 Human HLA
31 23 23.5 12 2 AAR67090 Taste mod
32 23 23.5 12 6 ABUI1918 Human HGP
33 23 23.5 14 4 AAB88073 CD66 pept
34 23 23.5 14 4 AAB88018 CD66 pept
35 23 23.5 14 4 AAB88099 CD66 pept
36 23 23.5 14 4 AAB88056 CD66 pept
37 23 23.5 14 5 ABP67248 Human CD6
38 23 23.5 14 6 ABUI1935 Human HGP
39 23 23.5 14 6 ABUI1953 Human HGP
40 23 23.5 15 6 ABUI1944 Human HGP
41 23 23.5 15 6 ABUI1956 Human HGP
42 23 23.5 15 6 ABP83373 G protein
43 23 23.5 15 7 ABM74254 DNA clone
44 23 23.5 18 5 AAU89150 Insulin/i
45 23 23.5 18 6 ADA03975 IGF-1R re

CC or nucleic acids encoding such polypeptides are useful in a composition
CC or as a part of the treatment regimen for the prevention of amelioration
CC of Streptococcus pneumoniae infection. The composition of the invention
CC is non-pathogenic and substantially free from any infectious impurities.
CC The immunogenic compositions can be compounded with fewer components to
CC elicit protection comparable to previously used agents. The present
CC sequence is the amino acid sequence of a Neisseria meningitidis strain-
CC specific mature 2086 protein with a native leader sequence which can be
CC used to create the composition of the invention.

XX SQ Sequence 260 AA;

Query Match 25.5%; Score 25; DB 7; Length 260;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXG 56
||| |||
Db 48 NGTLTSAQAQAEKTYGNGDSLNTG 71

RESULT 2
ADE44818
ID ADE44818 standard; protein; 260 AA.
XX ADE44818;
AC
XX 29-JAN-2004 (first entry)
XX
DE Neisseria meningitidis ORF2086 protein SeqID252.
XX
KW ORF2086; Neisseria meningitidis serogroup B infection; antibacterial;
KW antiinflammatory; immune response; bacterial meningitis;
KW Streptococcus pneumoniae infection; non-pathogenic;
KW immunogenic composition; 2086 protein.
XX Neisseria meningitidis.
OS
XX WO2003063766-A2.
PN
XX
PD 07-AUG-2003.
XX
PF 11-OCT-2002; 2002WO-US032369.
XX
PR 11-OCT-2001; 2001US-0328101P.
PR 30-AUG-2002; 2002US-0406934P.
XX
XX (AMHP) WYETH HOLDINGS CORP.
PA
XX Zlotnick GW, Fletcher LD, Farley J, Bernfield LA, Zagursky RJ;
PI Metcalf BJ;
PI
XX WPI; 2003-663416/62.
DR N-PSDB; ADE44817.
XX
PT Composition comprising crossreactive immunogenic antigen encoded by open
PT reading frame 2086 of Neisseria sp., that provides immunogenicity against
PT meningitis, or its immunogenic portion or biological equivalent.
XX
PS Claim 60; SEQ ID NO 252; 480pp; English.
XX
XX This invention relates to a novel composition which comprises at least
CC one protein (or fragment of) encoded by an open reading frame (ORF) of a
CC Neisseria sp. (ORF2086), where the ORF encoding a crossreactive
CC immunogenic antigen provides immunogenicity against infection by
CC Neisseria meningitidis serogroup B in a subject. The composition of the
CC invention may have antibacterial or antiinflammatory activity through the
CC induction of the immune response. The invention may be useful for the
CC treatment of bacterial meningitis in a mammal. One or more polypeptides
CC or nucleic acids encoding such polypeptides are useful in a composition
CC or as a part of the treatment regimen for the prevention of amelioration
CC of Streptococcus pneumoniae infection. The composition of the invention
CC is non-pathogenic and substantially free from any infectious impurities.

CC The immunogenic compositions can be compounded with fewer components to
CC elicit protection comparable to previously used agents. The present
CC sequence is the amino acid sequence of a Neisseria meningitidis strain-
CC specific mature 2086 protein which can be used to create the composition
CC of the invention.

XX SQ Sequence 260 AA;

Query Match 25.5%; Score 25; DB 7; Length 260;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXG 56
||| |||
Db 48 NGTLTSAQAQAEKTYGNGDSLNTG 71

RESULT 3
ADE44816
ID ADE44816 standard; protein; 261 AA.
XX ADE44816;
AC
XX 29-JAN-2004 (first entry)
XX
DE Neisseria meningitidis ORF2086 protein SeqID250.
XX
KW ORF2086; Neisseria meningitidis serogroup B infection; antibacterial;
KW antiinflammatory; immune response; bacterial meningitis;
KW Streptococcus pneumoniae infection; non-pathogenic;
KW immunogenic composition; 2086 protein; P4 leader.
XX Neisseria meningitidis.
OS
XX Haemophilus influenzae.
XX WO2003063766-A2.
PN
XX
PD 07-AUG-2003.
XX
PF 11-OCT-2002; 2002WO-US032369.
XX
PR 11-OCT-2001; 2001US-0328101P.
PR 30-AUG-2002; 2002US-0406934P.
XX
XX (AMHP) WYETH HOLDINGS CORP.
PA
XX Zlotnick GW, Fletcher LD, Farley J, Bernfield LA, Zagursky RJ;
PI Metcalf BJ;
PI
XX WPI; 2003-663416/62.
DR N-PSDB; ADE44815.
XX
PT Composition comprising crossreactive immunogenic antigen encoded by open
PT reading frame 2086 of Neisseria sp., that provides immunogenicity against
PT meningitis, or its immunogenic portion or biological equivalent.
XX
PS Claim 60; SEQ ID NO 250; 480pp; English.
XX
XX This invention relates to a novel composition which comprises at least
CC one protein (or fragment of) encoded by an open reading frame (ORF) of a
CC Neisseria sp. (ORF2086), where the ORF encoding a crossreactive
CC immunogenic antigen provides immunogenicity against infection by
CC Neisseria meningitidis serogroup B in a subject. The composition of the
CC invention may have antibacterial or antiinflammatory activity through the
CC induction of the immune response. The invention may be useful for the
CC treatment of bacterial meningitis in a mammal. One or more polypeptides
CC or nucleic acids encoding such polypeptides are useful in a composition
CC or as a part of the treatment regimen for the prevention of amelioration
CC of Streptococcus pneumoniae infection. The composition of the invention
CC is non-pathogenic and substantially free from any infectious impurities.
CC The immunogenic compositions can be compounded with fewer components to
CC elicit protection comparable to previously used agents. The present
CC sequence is the amino acid sequence of a Neisseria meningitidis strain-

CC specific mature 2086 protein with a p4 leader sequence (derived from the
 CC Haemophilus influenzae P4 protein) which can be used to create the
 CC composition of the invention.

XX SQ Sequence 261 AA;

Query Match 25.5%; Score 25; DB 7; Length 261;
 Best Local Similarity 25.0%; Pred. No. 1.5e+03;
 Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
 Db 49 NGTILSAQAQKTYGNGDSLNTG 72

RESULT 4

ADP70329
 ID ADP70329 standard; protein; 274 AA.

AC ADP70329;

DT 26-AUG-2004 (first entry)

XX Neisseria meningitidis NMB1870 protein, SEQ ID 128.

DE Antibacterial; Antiinflammatory; Vaccine; meningococcal protein; NMB1870;
 KW antigen; bactericide; meningitis; bacteraemia; Neisserial infection;
 KW lipoprotein.

XX Neisseria meningitidis; strain ISS1113.

OS WO2004048404-A2.

PN 10-JUN-2004.

PD 21-NOV-2003; 2003WO-IB006320.

PF 22-NOV-2002; 2002GB-00027346.

PR (CHIR) CHIRON SRL.

PA Comanducci M, Pizza M;

PI WPI; 2004-468293/44.

DR Composition for producing medicament for preventing Neisserial infection
 PT in mammal, comprising multiple variant of meningococcal protein NMB1870,
 PT as antigen.

PS Claim 11; SEQ ID NO 128; 77pp; English.

XX The present invention relates to composition (I) comprising at least two
 CC multiple variants of meningococcal protein NMB1870 (ADP70202-ADP70246,
 CC ADP70278, ADP70280-ADP70286, ADP70288-ADP70295, ADP70324-ADP70343) as
 CC antigen. (I) elicit a bactericidal response effective against each of
 CC serogroup B Neisseria meningitidis strains MC58, 961-5945 and M1239, and
 CC also elicit an antibody response which is bactericidal against N.
 CC meningitidis strains in at least 2 of hypervirulent lineages ET-37, ET-5,
 CC cluster A4, lineage 3, subgroup I, subgroup III, and subgroup IV-1. In
 CC NMB1870 is a lipoprotein. In (I) at least one of the proteins does not
 CC include the amino acid sequence ADP70271 or ADP70272 within 10 amino
 CC acids of its N-terminus or at least one of the proteins does not include
 CC the amino acid sequences ADP70273 within 10 amino acids of its N-terminus
 CC or at least one of the proteins includes the amino acid sequence
 CC ADP70274. ADP70271-ADP70274 are sequence motifs for retention or omission
 CC from the NMB1870 proteins. (I) or (II) is efficiently elicits systemic
 CC and/or mucosal immunity and thereby prevents/treat meningitis and
 CC bacteraemia. (I) is useful as medicament and for raising an antibody
 CC response in a mammal for protecting a mammal against a Neisserial
 CC infection.

XX SQ Sequence 274 AA;

Query Match 25.5%; Score 25; DB 8; Length 274;
 Best Local Similarity 25.0%; Pred. No. 1.6e+03;
 Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
 Db 62 NGTILSAQAQKTYGNGDSLNTG 85

RESULT 5

AAR48993
 ID AAR48993 standard; protein; 1026 AA.

AC AAR48993;

DT 16-OCT-2003 (revised)

DT 14-SEP-1994 (first entry)

DE rsaA S-lyase protein.

KW C. crescentus; rsaA; paracrystalline; S-layer; protein; heterologous;
 KW cellulase; xylase; metallothionein; restriction site; reading frame;
 KW fusion protein; bioreactor; toxic metal; sewage; waste water;
 KW wood pulp suspension; cell surface; vaccine; fish.

XX Caulobacter vibrioides.

XX CA2090549-A.

PN 10-DEC-1993.

PP 26-FEB-1993; 93CA-02090549.

PR 09-JUN-1992; 92US-00895367.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX Smit J, Bingle WH;

XX WPI; 1994-066249/09.

DR N-PSDE; AAQ57972.

XX Prodn. of heterologous polypeptides in bacteria, partic. Caulobacter - by
 PT expression of a fusion prod. of the polypeptide sequence and a bacterial
 PT S-layer protein gene.

PS Claim 17; Fig 6; 27pp; English.

XX This sequence is encoded by the C. crescentus rsaA gene and represents
 CC the paracrystalline S-layer protein. The rsaA gene was used in the
 CC production of the heterologous protein of the invention. The heterologous
 CC protein is produced by cloning a polypeptide coding sequence, eg.
 CC cellulase, xylase or a metallothionein, into a restriction site within
 CC the rsaA gene which preserves the rsaA reading frame and expressing the
 CC fusion sequence in Caulobacter. This S-layer protein bacterial system
 CC can be used in bioreactors, eg. to bind toxic metals in sewage waste
 CC water etc. or for the treatment of wood pulp suspensions. The system can
 CC be used to produce heterologous proteins at the cell surface for use in
 CC vaccines, partic. fish vaccines. The S-layer protein is synthesised in
 CC large quantities and has a general repetitive sequence, permitting the
 CC synthesis of large amounts of heterologous protein as a fusion product
 CC and presentation at the cell surface. (Updated on 16-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 1026 AA;

Query Match 25.5%; Score 25; DB 2; Length 1026;
 Best Local Similarity 25.0%; Pred. No. 6.2e+03;
 Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
 Db 267 NDTFVAGEVAGATLTIVGDTLSGG 290

```

RESULT 6
ID AAR94014 standard; protein; 1026 AA.
XX
AC AAR94014;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 21-MAY-1996 (first entry)
XX
XX Caulobacter S-layer protein.
DE
XX
XX S-layer; rsaA gene; surface protein; biofilm; vaccine.
XX
XX Caulobacter vibrioides; strain CB15.
OS
XX
XX Key Location/Qualifiers
FH 860..905
FT Region
FT /note= "glycine-aspartic acid repeat region"
XX
XX US5500353-A.
XX
XX 19-MAR-1996.
XX
XX 09-FEB-1994; 94US-00194290.
XX
XX 09-JUN-1992; 92US-00895367.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Smit J, Bingle WH;
PI
XX WPI; 1996-171046/17.
DR N-PSDB; AAT17717.
XX
XX Expression and presentation of polypeptide heterologous to Caulobacter S-
PT layer protein to bacterium's environment - by cloning in-frame into
PT Caulobacter rsaA gene, and expressing as fusion prod. with S-layer
PT protein.
XX
XX Example 2; Fig 6a-c; 22pp; English.
XX
XX The Caulobacter crescentus CB15 paracrystalline S-layer protein
CC (AAR17717) is encoded by the rsaA gene (AAT17717). It can be used as a
CC fusion partner with polypeptides of interest, allowing presentation of
CC the polypeptide on the surface of the Caulobacter cells. The bacterium is
CC cultured as a biofilm in a bioreactor or may be used to present an
CC antigenic epitope (see AAR94016 and AAR94018) to the environment e.g. for
CC use as a vaccine. (Updated on 25-MAR-2003 to correct PF field.) (Updated
CC on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 1026 AA;
QY Query Match 25.5%; Score 25; DB 2; Length 1026;
Best Local Similarity 25.0%; Pred. No. 6.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Db 267 NDTFVAGEVAGAAATLVGDTLGG 290
QY 33 NGTFXXXXXXXGDXXXG 56
Db 267 NDTFVAGEVAGAAATLVGDTLGG 290
RESULT 7
ID AAW37490 standard; protein; 1026 AA.
XX
XX AAW37490;
XX
XX 17-OCT-2003 (revised)
DT 20-APR-1998 (first entry)
XX
XX

```

```

DE Caulobacter crescentus S-layer rsaA protein.
XX
XX S-layer; rsaA gene; Caulobacter; vaccine; antigenic; ligand; enzyme;
XX metallothionein; heavy metal; water; sewage; xylanase; cellulase;
XX wood pulping.
XX
XX Caulobacter vibrioides.
OS
XX
XX WO9734000-A1.
PN
XX
XX 18-SEP-1997.
PD
XX
XX 10-MAR-1997; 97WO-CA000167.
PF
XX
XX 12-MAR-1996; 96US-00614377.
PR
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
PA
XX
XX Smit J, Bingle WH, Nomellini JF;
PI
XX WPI; 1997-470880/43.
DR N-PSDB; AAV01866.
XX
XX New DNA containing sequence for C-terminal region of Caulobacter S-layer
PT protein - expressed as fusion proteins containing antigenic peptides in
PT Caulobacter, useful as live vaccines.
XX
XX Example 2; Fig 6; 58pp; English.
XX
XX The present sequence represents a Caulobacter S-layer protein used in an
CC example of the present invention. A new DNA construct has been developed
CC which contains at least one restriction site for insertion of DNA
CC upstream of DNA encoding a C-terminal region of at least the last 82
CC amino acids (aa) of Caulobacter S-layer protein. Caulobacter containing a
CC the DNA constructs above additionally containing a sequence encoding a
CC heterologous polypeptide, are particularly useful in live vaccines (where
CC the heterologous polypeptide is an antigen). They can also be used for
CC production of e.g. ligands, enzymes or other proteins, e.g.
CC metallothioneins to remove heavy metals from water or sewage, or xylanase
CC or cellulase for use in wood pulping. All known Caulobacter strains are
CC harmless, and stable in outdoor environments, including water (so
CC suitable for vaccinating fish) or soil. They are well suited for growing
CC in biofilm reactors and produce S-layer proteins, which is an ideal
CC system for presentation of antigens, at high level. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX
XX Sequence 1026 AA;
QY Query Match 25.5%; Score 25; DB 2; Length 1026;
Best Local Similarity 25.0%; Pred. No. 6.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 33 NGTFXXXXXXXGDXXXG 56
Db 267 NDTFVAGEVAGAAATLVGDTLGG 290
RESULT 8
ID AAY44757 standard; protein; 1026 AA.
XX
XX AAY44757;
XX
XX 12-SEP-2003 (revised)
DT 04-MAY-2000 (first entry)
XX
XX Caulobacter crescentus surface layer protein.
DE
XX
XX Surface layer protein; S-layer secretion signal; antibiotic; vaccine;
XX recombinant fusion protein cleavage; enzyme; protein polymer;
XX antibacterial enzyme; foodstuff.
XX
XX Caulobacter vibrioides.
OS

```


XX Claim 1; SEQ ID NO 266; 244pp; English.

XX This invention relates to a novel isolated peptide which comprises at

CC least 1, 2, 3, 4, 5, 6 or 7 epitopes or analogues of the epitopes given

CC in the specification. The peptide comprises HLA-A1, A2, A3, A24, B7 or

CC B44 tumour associated antigen peptide. The invention may be useful for

CC the production of compounds with a cytostatic activity or for the

CC production of a vaccine. The peptide is useful in preparing a composition

CC diagnosing or treating tumour associated antigen-related disease. The

CC present sequence is that of an epitope peptide for use in the peptide of

CC the invention.

XX Sequence 8 AA;

SQ

Query Match 23.5%; Score 23; DB 8; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36

Db 5 NGTF 8

RESULT 11

ADD23187

ID ADD23187 standard; peptide; 9 AA.

XX

AC ADD23187;

XX

DT 15-JAN-2004 (first entry)

XX

DE Breast cancer membrane protein (BCMP) peptide SEQ ID NO:617.

XX

DE Breast cancer; screening; diagnosis; breast cancer therapy;

KW breast cancer membrane protein; BCMP; cytostatic; vaccine; human.

KW

XX Homo sapiens.

OS

XX WO2003087831-A2.

PN

XX 23-OCT-2003.

PD

XX 10-APR-2003; 2003WO-GB001559.

PF

XX 11-APR-2002; 2002GB-00008331.

PR

XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

PA

XX Hudson LJ, Stamps AC, Terrett JA;

PI

XX WPI; 2003-845381/78.

DR

XX Screening, diagnosing and/or treating breast cancer by detecting a change

PT in expression or activity of a breast cancer membrane protein (BCMP)

PT polypeptide or encoding nucleic acid molecule.

XX

XX Example; SEQ ID NO 617; 81pp; English.

PS

XX The present invention describes a method of screening for and/or

CC diagnosing breast cancer in a subject, and/or monitoring the

CC effectiveness of breast cancer therapy. The method comprises detecting

CC and/or quantifying in a biological sample obtained from the subject a

CC breast cancer membrane protein (BCMP) polypeptide and a nucleic acid

CC molecule. Also described: (1) an antibody, its functionally-active

CC fragment, derivative or analogue, that specifically binds to one or more

CC of the BCMP polypeptide; (2) a diagnostic kit comprising a capture

CC reagent specific for an BCMP polypeptide, reagents and instructions for

CC use; (3) a method for screening for anti-breast cancer agents that

CC interact with the BCMP polypeptide, comprising contacting the polypeptide

CC with a candidate agent, and determining whether or not the candidate

CC agent interacts with the polypeptide; (4) a method for screening for anti

CC -breast cancer agents that modulate the expression or activity of an BCMP

CC

polypeptide or the nucleic acid molecule cited above, comprising

CC comparing the expression or activity of the polypeptide or nucleic acid

CC molecule, in the presence and absence of a candidate agent or in the

CC presence of a control agent, and determining whether the candidate agent

CC causes the expression or activity of the polypeptide or nucleic acid

CC molecule to change; and (5) an agent identified by the method of (3) or

CC (4), which interacts with the polypeptide or causes the expression or

CC activity of the polypeptide, or the expression of the nucleic acid

CC molecule to change. BCMPs have cytostatic activities, and can be used in

CC vaccines. The BCMP polypeptide, nucleic acid molecule, antibody, agent or

CC their derivatives, are useful in the manufacture of a medicament for the

CC treatment of breast cancer, where the composition is a vaccine. The

CC present sequence represents a BCMP peptide which is used in the

CC exemplification of the present invention.

XX

SQ Sequence 9 AA;

Query Match 23.5%; Score 23; DB 7; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36

Db 4 NGTF 7

RESULT 12

ADM72935

ID ADM72935 standard; peptide; 9 AA.

XX

AC ADM72935;

XX

DT 03-JUN-2004 (first entry)

XX

DE Human CEA epitope SEQ ID NO:194.

XX

XX epitope; epitope cluster; virucide; cytostatic; vaccine; viral infection;

KW cancer; tumour; human; CEA.

KW

XX Homo sapiens.

OS

OS Synthetic.

XX

PN WO2004022709-A2.

XX

PD 18-MAR-2004.

XX

PF 05-SEP-2003; 2003WO-US027706.

XX

PR 06-SEP-2002; 2002US-0409123P.

XX

XX (MANN-) MANNKIND CORP.

PA

XX Simard JLL, Diamond DC, Liu L, Liu Z;

PI

XX WPI; 2004-315564/29.

DR

XX New polypeptides and encoding nucleic acids that are useful epitopes of

PT target-associated antigens, useful for diagnosing and/or treating viral

PT infections, cancers and tumors.

XX

PS Claim 1; SEQ ID NO 194; 357pp; English.

XX

XX The present invention describes a polypeptide (I) comprising a component

CC selected from: (a) a polypeptide epitope having any of the 503 fully

CC defined sequences of 8-33 amino acids (SEQ ID NO:108-610); (b) an epitope

CC cluster comprising the polypeptide of (a); (c) a polypeptide having

CC substantial similarity to (a) or (b); (d) a polypeptide having functional

CC similarity to any of (a)-(c); or (e) a nucleic acid encoding the

CC polypeptide of (a)-(d). (I) has virucide and cytostatic activities, and

CC can be used in vaccines. The methods and compositions of the present

CC invention are useful for the diagnosis and/or treatment of viral

CC infections, cancers and tumors. The present sequence is used in the

CC exemplification of the present invention.

```

XX SQ Sequence 9 AA;
Query Match 23.5%; Score 23; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
   ||||
Db 3 NGTF 6

RESULT 13
AAV46550
ID AAY46550 standard; peptide; 10 AA.
XX AC AAY46550;
XX DT 01-DEC-1999 (first entry)
XX DE Immunogenic peptide having a human leukocyte antigen binding motif #1161.
XX KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX KW immune response; T cell activation; major histocompatibility complex;
XX KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX KW vaccine; immunisation.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9945954-A1.
XX PD 16-SEP-1999.
XX PF 13-MAR-1998; 98WO-US005039.
XX PR 13-MAR-1998; 98WO-US005039.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI; 1999-551214/46.
XX PT New immunogenic peptides with HLA binding motif, useful in treatment and
XX diagnosis of cancers and viral diseases.
XX PS Claim 1; Page 76; 150pp; English.
XX CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
XX having a human major histocompatibility complex (MHC) Class I (also known
XX as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
XX can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
XX or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
XX the antigen from which the peptide is derived. Cytotoxic T lymphocytes
XX (CTLs) which destroy antigen-bearing cells are normally induced by an
XX antigen in the form of a peptide fragment bound to a HLA molecule, rather
XX than the intact foreign antigen itself, and are particularly important in
XX tumour rejection and in fighting viral infections. The peptides are
XX therefore useful therapeutically to treat or prevent viral infections and
XX cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
XX and C, AIDS, and renal carcinoma. They can be administered as vaccines to
XX elicit an immune response in individuals susceptible or otherwise at risk
XX of viral infection or cancer, or used to treat chronic or acute
XX conditions. They are also useful diagnostically, and can be used to
XX induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
XX the peptide e.g. to produce CTLs ex vivo for infusion back into a
XX patient. The polynucleotides encoding the immunogenic peptides are also
XX useful therapeutically and for immunisation as above
XX SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
   ||||
Db 7 NGTF 10

RESULT 14
AAV00829
ID AAY00829 standard; peptide; 10 AA.
XX AC AAY00829;
XX DT 19-MAY-1999 (first entry)
XX DE HLA-A24 antigenic peptide CB-202.
XX KW HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
XX KW cancer; therapy; diagnosis.
XX OS Synthetic.
XX PN WO9903972-A1.
XX PD 28-JAN-1999.
XX PF 13-JUL-1998; 98WO-JF003143.
XX PR 15-JUL-1997; 97JP-00203900.
XX PR 15-JUL-1997; 97JP-00203917.
XX PR 12-JAN-1998; 98JP-00014736.
XX PA (TAKI) TAKARA SHUZO CO LTD.
XX PI Takesako K, Nukaya I, Yasumoto M, Iwasaki T, Ideno M, Akiyoshi T;
XX PI Fujie T, Tanaka F, Kato I;
XX WPI; 1999-132220/11.
XX PT Cytotoxic T-lymphocytes recognising cells expressing HLA-A24/antigenic
XX peptide complex - and inducing agents for them consisting of such antigen
XX -presenting cells.
XX PS Example 9; Page 70; 88pp; Japanese.
XX CC This sequence represents an HLA-A24 antigenic peptide sequence used in
XX the cytotoxic T-lymphocytes (CTL) of the invention. The CTL recognise
XX cells expressing on their surface a complex of an HLA-A24 molecule with
XX an antigenic peptide. The antigenic peptide is an HLA-A24 restrained
XX antigenic peptide. The CTLs can be used in the treatment and diagnosis of
XX cancers
XX SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
   ||||
Db 7 NGTF 10

RESULT 15
AAV00830
ID AAY00830 standard; peptide; 10 AA.
XX AC AAY00830;
XX DT 19-MAY-1999 (first entry)
XX
```

```

DE HLA-A24 antigenic peptide CE-203.
XX
XX HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
XX cancer; therapy; diagnosis.
XX
XX Synthetic.
XX
XX WO9903972-A1.
XX
XX 28-JAN-1999.
XX
XX 13-JUL-1998; 98WO-JP003143.
XX
XX 15-JUL-1997; 97JP-00203900.
XX 15-JUL-1997; 97JP-00203917.
XX 12-JAN-1998; 98JP-00014736.
XX
XX (TAKI ) TAKARA SHUZO CO LTD.
XX
XX Takesako K, Nukaya I, Yasumoto M, Iwasaki T, Ideno M, Akiyoshi T;
XX Fujie T, Tanaka F, Kato I;
XX WPI; 1999-132220/11.
XX
XX Cytotoxic T-lymphocytes recognising cells expressing HLA-A24/antigenic
XX peptide complex - and inducing agents for them consisting of such antigen
XX -presenting cells.
XX
XX Example 9; Page 76; 88pp; Japanese.
XX
XX This sequence represents an HLA-A24 antigenic peptide sequence used in
XX the cytotoxic T-lymphocytes (CTL) of the invention. The CTL recognise
XX cells expressing on their surface a complex of an HLA-A24 molecule with
XX an antigenic peptide. The antigenic peptide is an HLA-A24 restrained
XX antigenic peptide. The CTLs can be used in the treatment and diagnosis of
XX cancers
XX
XX Sequence 10 AA;
SQ
Query Match 23.5%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 NGTF 36
Db |||||
7 NGTF 10
XX
XX RESULT 16
XX AAY00835
XX ID AAY00835 standard; peptide; 10 AA.
XX AC AAY00835;
XX
XX 19-MAY-1999 (first entry)
XX
XX HLA-A24 antigenic peptide CE-208.
XX
XX HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
XX cancer; therapy; diagnosis.
XX
XX Synthetic.
XX
XX WO9903972-A1.
XX
XX 28-JAN-1999.
XX
XX 13-JUL-1998; 98WO-JP003143.
XX
XX 15-JUL-1997; 97JP-00203900.
XX 15-JUL-1997; 97JP-00203917.
XX 12-JAN-1998; 98JP-00014736.
XX
XX (TAKI ) TAKARA SHUZO CO LTD.
XX
XX Takesako K, Nukaya I, Yasumoto M, Iwasaki T, Ideno M, Akiyoshi T;
XX Fujie T, Tanaka F, Kato I;
XX WPI; 1999-132220/11.
XX
XX Cytotoxic T-lymphocytes recognising cells expressing HLA-A24/antigenic
XX peptide complex - and inducing agents for them consisting of such antigen
XX -presenting cells.
XX
XX Example 9; Page 76; 88pp; Japanese.
XX
XX This sequence represents an HLA-A24 antigenic peptide sequence used in
XX the cytotoxic T-lymphocytes (CTL) of the invention. The CTL recognise
XX cells expressing on their surface a complex of an HLA-A24 molecule with
XX an antigenic peptide. The antigenic peptide is an HLA-A24 restrained
XX antigenic peptide. The CTLs can be used in the treatment and diagnosis of
XX cancers
XX
XX Sequence 10 AA;
SQ
Query Match 23.5%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 NGTF 36
Db |||||
7 NGTF 10
XX
XX RESULT 16
XX AAY00833
XX ID AAY00833 standard; peptide; 10 AA.
XX
XX AC AAY00833;
XX
XX 19-MAY-1999 (first entry)
XX
XX HLA-A24 antigenic peptide CE-206.
XX
XX HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
XX cancer; therapy; diagnosis.
XX
XX Synthetic.
XX
XX WO9903972-A1.
XX
XX 28-JAN-1999.
XX
XX 13-JUL-1998; 98WO-JP003143.
XX
XX 15-JUL-1997; 97JP-00203900.
XX 15-JUL-1997; 97JP-00203917.
XX 12-JAN-1998; 98JP-00014736.
XX
XX
XX

```

CC antigenic peptide. The CTLs can be used in the treatment and diagnosis of
 CC cancers
 XX
 SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 7 NGTF 10

RESULT 18
 AAY00809
 ID AAY00809 standard; peptide; 10 AA.
 XX
 AC AAY00809;
 XX
 DT 19-MAY-1999 (first entry)
 XX
 DE HLA-A24 antigenic peptide, CEA peptide CE-2.
 XX
 KW HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
 KW cancer; therapy; diagnosis.
 XX
 OS Synthetic.
 XX
 PN WO9903972-A1.
 XX
 PD 28-JAN-1999.
 XX
 PF 13-JUL-1998; 98WO-JP003143.
 XX
 PR 15-JUL-1997; 97JP-00203900.
 PR 15-JUL-1997; 97JP-00203917.
 PR 12-JAN-1998; 98JP-00014736.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Takesako K, Nukaya I, Yasumoto M, Iwasaki T, Ideno M, Akiyoshi T;
 PI Fujie T, Tanaka F, Kato I;
 XX
 DR WPI; 1999-132220/11.
 XX
 CC Cytotoxic T-lymphocytes recognising cells expressing HLA-A24/antigenic
 PT peptide complex - and inducing agents for them consisting of such antigen
 PT -presenting cells.
 XX
 PS Claim 1; Page 62; 88pp; Japanese.
 CC
 CC This sequence represents an HLA-A24 antigenic peptide sequence used in
 CC the cytotoxic T-lymphocytes (CTL) of the invention. The CTL recognise
 CC cells expressing on their surface a complex of an HLA-A24 molecule with
 CC an antigenic peptide. The antigenic peptide is an HLA-A24 restrained
 CC antigenic peptide. The CTLs can be used in the treatment and diagnosis of
 CC cancers
 XX
 SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 7 NGTF 10

RESULT 19
 AAY00831
 ID AAY00831 standard; peptide; 10 AA.

XX AAY00831;
 AC
 XX 19-MAY-1999 (first entry)
 DT
 XX
 DE HLA-A24 antigenic peptide CE-204.
 XX
 KW HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
 KW cancer; therapy; diagnosis.
 XX
 OS Synthetic.
 OS
 XX
 PN WO9903972-A1.
 XX
 PD 28-JAN-1999.
 XX
 PF 13-JUL-1998; 98WO-JP003143.
 XX
 PR 15-JUL-1997; 97JP-00203900.
 PR 15-JUL-1997; 97JP-00203917.
 PR 12-JAN-1998; 98JP-00014736.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Takesako K, Nukaya I, Yasumoto M, Iwasaki T, Ideno M, Akiyoshi T;
 PI Fujie T, Tanaka F, Kato I;
 XX
 DR WPI; 1999-132220/11.
 XX
 CC Cytotoxic T-lymphocytes recognising cells expressing HLA-A24/antigenic
 PT peptide complex - and inducing agents for them consisting of such antigen
 PT -presenting cells.
 XX
 PS Example 9; Page 76; 88pp; Japanese.
 CC
 CC This sequence represents an HLA-A24 antigenic peptide sequence used in
 CC the cytotoxic T-lymphocytes (CTL) of the invention. The CTL recognise
 CC cells expressing on their surface a complex of an HLA-A24 molecule with
 CC an antigenic peptide. The antigenic peptide is an HLA-A24 restrained
 CC antigenic peptide. The CTLs can be used in the treatment and diagnosis of
 CC cancers
 XX
 SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 7 NGTF 10

RESULT 20
 AAY00828
 ID AAY00828 standard; peptide; 10 AA.
 XX
 AC AAY00828;
 XX
 DT 19-MAY-1999 (first entry)
 XX
 DE HLA-A24 antigenic peptide CE-201.
 XX
 KW HLA-A24; antigenic peptide; cytotoxic T-lymphocyte; CTL; HLA-A24 complex;
 KW cancer; therapy; diagnosis.
 XX
 OS Synthetic.
 OS
 XX
 PN WO9903972-A1.
 XX
 PD 28-JAN-1999.
 XX
 PF 13-JUL-1998; 98WO-JP003143.

RESULT 23
 ID AAU82066 standard; peptide; 10 AA.
 XX AC ADF83690;
 XX AC ADF83690;
 DT 26-FEB-2004 (first entry)
 XX DE CBA antigenic peptide C3 associated with cancer.
 XX DE T-cell binding ligand; TCBL; peptide G'; human MHC class II beta chain;
 KW peptide J; human beta-2-microglobulin; HIV-1; TCBL peptide construct;
 KW immunological disorder; immune response; human immunodeficiency virus;
 KW herpes simplex virus infection; HSV; malaria; tuberculosis; cancer; CEA;
 KW acquired immunodeficiency syndrome; AIDS; allergy; autoimmune disease;
 KW autoimmune myocarditis; cytostatic; antiinflammatory.
 XX OS Synthetic.
 XX PN WO200189286-A2.
 XX PD 29-NOV-2001.
 XX PF 24-MAY-2001; 2001WO-US016793.
 XX PR 24-MAY-2000; 2000US-0206548P.
 XX PA (CELS-) CEL-SCI CORP.
 XX PI Zimmerman DS, Sarin PS;
 XX PT WPI; 2002-083037/11.
 XX DR New T cell binding ligand peptide for treating immunological disorders
 XX PT such as herpes simplex virus, tuberculosis, cancers, acquired
 XX PT immunodeficiency syndrome and allergies.
 XX PS Disclosure; Page 27; 110pp; English.
 XX SS The present invention relates to novel T-cell binding ligand (TCBL)
 CC peptides (e.g. peptide G' (modified human MHC class II beta chain peptide
 CC G, peptide J' (human beta-2-microglobulin peptide) and HIV-1 peptides) and
 CC TCBL peptide constructs for treating immunological disorders. The peptide
 CC constructs are useful for eliciting a cellular immune response in a human
 CC patient. The method comprises administering the peptide construct to the
 CC patient preferably in combination with an immune response adjuvant. The
 CC peptide constructs in the form of conjugated peptides are useful for
 CC eliciting a cellular immune response in a patient exposed to or at risk
 CC for exposure to the human immunodeficiency virus (HIV). The TCBL peptides
 CC are useful for treating a patient suffering from an immunological
 CC disorder such as herpes simplex virus (HSV) infection, malaria,
 CC tuberculosis, cancers, acquired immunodeficiency syndrome (AIDS),
 CC allergies, autoimmune diseases (e.g. arthritis, Graves disease, multiple
 CC sclerosis (MS), autoimmune myocarditis, diabetes and lupus) by
 CC administering a peptide construct comprising a TCBL peptide bonded to an
 CC antigenic peptide associated with the disorder. Unlike prior art peptide
 CC conjugates, a modified version of peptide G has long range stabilisation
 CC and also enhances the immune response. AAU82019-AAU82114 represent T-cell
 CC specific binding ligand peptides, peptide constructs or peptides used in
 CC their construction
 XX SQ Sequence 10 AA;
 Query Match 23.5%; Score 23; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 NGTF 36
 DB 7 NGTF 10
 RESULT 24
 ID ADF83690 standard; peptide; 10 AA.
 XX AC ADF83690;
 XX AC ADF83690;
 DT 26-FEB-2004 (first entry)
 XX DE HLA-A24 related human peptide, SEQ ID No 20.
 XX DE non-human transgenic mammal; HLA-A24; cytotoxic T cell; CTL; tumour;
 KW viral infection; antigen; chimeric; cytostatic; virucide; human.
 XX OS Homo sapiens.
 XX PN WO200247474-A1.
 XX PD 20-JUN-2002.
 XX PF 12-DEC-2001; 2001WO-JP010885.
 XX PR 13-DEC-2000; 2000JP-00378556.
 XX PR 06-SEP-2001; 2001JP-00269746.
 XX PA (SUMU) SUMITOMO PHARM CO LTD.
 XX PI Gotoh M;
 XX PT WPI; 2003-067332/06.
 XX DR Non-human transgenic mammals expressing HLA-A24 with cytotoxic T cells
 XX PT induced by HLA-A24-restricted antigenic stimulation, useful in screening
 XX PT preventives or remedies for tumor or viral infections.
 XX PS Example 9; SEQ ID NO 20; 92pp; Japanese.
 XX SS The invention further relates to a novel non-human transgenic mammal
 CC transfected with an HLA-A24 gene in which cytotoxic T cells (CTL) are
 CC induced by HLA-A24-restricted antigenic stimulation. The invention
 CC further relates to a novel method of screening a remedy or a preventative
 CC for a tumour or viral infection characterised by comprising administering
 CC a test substance to the above-described non-human transgenic mammal and
 CC then assaying and evaluating whether or not CTLs specific to the test
 CC substance are induced therein. The invention also includes: a PSA-origen
 CC HLA-A24-restricted tumour antigen peptide selected by this screening
 CC method; a chimeric gene useful in forming the above-described transgenic
 CC non-human mammal; host cells transformed by this chimeric gene; and
 CC utilisation thereof. The novel non-human transgenic mammal has cytostatic
 CC and virucide activity. This sequence represents a human HLA-A24 related
 CC peptide of the invention.
 XX SQ Sequence 10 AA;
 Query Match 23.5%; Score 23; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 NGTF 36
 DB 7 NGTF 10
 RESULT 25
 ID ADG38740 standard; peptide; 10 AA.
 XX AC ADG38740;
 XX AC ADG38740;
 DT 26-FEB-2004 (first entry)
 XX DE Human carcinoembryonic antigen-derived peptide #43.

KW major histocompatibility complex class I epitope; MHC class I epitope;
KW enhanced immunogenicity; heteroclitic analogue; immune response;
KW viral disease; cancer; human; carcinoembryonic antigen.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003087126-A2.
XX
PD 23-OCT-2003.
XX
XX 07-APR-2003; 2003WO-US010571.
PF
XX 05-APR-2002; 2002US-00116118.
PR
XX 26-SEP-2002; 2002US-0413471P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
XX Ishioka G, Fikes J, Tangri S, Sette A;
PI WPI; 2003-865348/80.
XX
DR Producing analogs of Major Histocompatibility Complex (MHC) class I
XX peptide epitopes, for treating viral diseases or cancer, comprise
PT producing an analog having (semi)conservative amino acid substitutions of
PR the MHC class I epitope.
XX
XX Example 14; SEQ ID NO 178; 244pp; English.
PS
XX The invention comprises a method for producing a polypeptide that
XX contains an analogue of a major histocompatibility complex (MHC) class I
CC peptide epitope, wherein the analogue has enhanced immunogenicity
CC compared to the MHC class I peptide epitope. The method of the invention
CC is useful for producing an analogue or heteroclitic analogue of an MHC
CC class I peptide epitope, and eliciting an immune response. The method is
CC useful for preparing MHC class I peptide epitope analogues that can be
CC used to treat viral diseases, cancer and other conditions characterized
CC by displayed antigens on target cells. The present amino acid sequence
CC represents a peptide of the invention which is derived from human
CC carcinoembryonic antigen.
XX
SQ Sequence 10 AA;
Query Match 23.5%; Score 23; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 NGTF 36
Db |||||
7 NGTF 10
RESULT 26
ADG38664
ID ADG38664 standard; peptide; 10 AA.
XX
AC ADG38664;
XX
XX 26-FEB-2004 (first entry)
DT
XX Human carcinoembryonic antigen-derived peptide #22.
DE
XX major histocompatibility complex class I epitope; MHC class I epitope;
KW enhanced immunogenicity; heteroclitic analogue; immune response;
KW viral disease; cancer; human; carcinoembryonic antigen.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO2003087126-A2.
PN
XX 23-OCT-2003.
PD
XX

PF 07-APR-2003; 2003WO-US010571.
XX
XX 05-APR-2002; 2002US-00116118.
PR
XX 26-SEP-2002; 2002US-0413471P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
XX Ishioka G, Fikes J, Tangri S, Sette A;
PI WPI; 2003-865348/80.
XX
DR Producing analogs of Major Histocompatibility Complex (MHC) class I
XX peptide epitopes, for treating viral diseases or cancer, comprise
PT producing an analog having (semi)conservative amino acid substitutions of
PR the MHC class I epitope.
XX
XX Example 14; SEQ ID NO 102; 244pp; English.
PS
XX The invention comprises a method for producing a polypeptide that
XX contains an analogue of a major histocompatibility complex (MHC) class I
CC peptide epitope, wherein the analogue has enhanced immunogenicity
CC compared to the MHC class I peptide epitope. The method of the invention
CC is useful for producing an analogue or heteroclitic analogue of an MHC
CC class I peptide epitope, and eliciting an immune response. The method is
CC useful for preparing MHC class I peptide epitope analogues that can be
CC used to treat viral diseases, cancer and other conditions characterized
CC by displayed antigens on target cells. The present amino acid sequence
CC represents a peptide of the invention which is derived from human
CC carcinoembryonic antigen.
XX
SQ Sequence 10 AA;
Query Match 23.5%; Score 23; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 NGTF 36
Db |||||
7 NGTF 10
RESULT 27
AAO24401
ID AAO24401 standard; peptide; 10 AA.
XX
AC AAO24401;
XX
XX 06-MAY-2004 (first entry)
DT
XX HLA-A24-restricted cancer antigen peptide related peptide #35.
DE
XX Human; mouse; HLA-A24-restricted cancer antigen; antigen; cancer;
KW tumour suppressor protein; cytostatic; WTI; vaccine.
XX
XX Synthetic.
OS
XX WO2003106682-A1.
PN
XX 24-DEC-2003.
PD
XX 12-JUN-2003; 2003WO-JP007463.
PF
XX 12-JUN-2002; 2002JP-00171518.
PR
XX 20-SEP-2002; 2002JP-00275572.
FR
XX (CHUS) CHUGAI SEIYAKU KK.
PA (SUMU) SUMITOMO PHARM CO LTD.
XX (SUGI/) SUGIYAMA H.
PA
XX Sugiyama H, Gotoh M, Takasu H;
PI WPI; 2004-090846/09.
XX
XX

PT Antigenic peptides derived from Wt1 which induce HLA-A24 restricted
 PT cytotoxic T-lymphocytes for production of cancer vaccine and treatment
 PT and prevention of cancer.
 XX Disclosure; Page 97; Opp; Japanese.
 CC The present invention relates to antigenic peptides derived from tumour
 CC suppressor protein Wt1 which induce HLA-A24 restricted cytotoxic T-
 CC lymphocytes. The peptides can be used in the preparation of cancer
 CC vaccine for treatment and prevention of cancer, including leukaemia,
 CC multiple myeloma, lymphoma, and cancer of the stomach, colon, breast,
 CC liver, ovary, skin, pancreas, prostate and womb. The present sequence is
 CC a polypeptide used in the exemplification of the invention
 XX
 SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 8; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 ||||
 Db 7 NGTF 10

RESULT 28
 ADM12345
 ID ADM12345 standard; peptide; 10 AA.
 XX
 AC ADM12345;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE MHC class I epitope of human carcinoembryonic antigen, 9-5-4.
 XX
 KW antigen presentation enhancing hybrid polypeptide; mammalian II-key;
 KW MHC class II: antibacterial; virucide; fungicide; antirheumatic;
 KW antihistaminic; neuroprotective; dermatological; immunosuppressive;
 KW antinflammatory; antidiabetic; antithyroid; immune;
 KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
 KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
 KW scleroderma; dermatomyositis; pemphigus.
 XX
 OS Homo sapiens.
 XX
 PN US2003235594-A1.
 XX
 PD 25-DEC-2003.
 XX
 PF 17-SEP-2002; 2002US-00245871.
 XX
 PR 14-SEP-1999; 99US-00396813.
 PR 17-JUL-2002; 2002US-00197000.
 XX
 PA (ANTI-) ANTIGEN EXPRESS INC.
 XX
 PI Humphreys R, Xu M;
 XX
 DR WPI; 2004-070554/07.
 XX
 XX Novel II-key/antigen presentation enhancing hybrid polypeptide, useful
 PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
 PT erythematosus and diabetes mellitus.
 XX
 PS Example 9; Page 34; 87pp; English.
 XX

CC The invention relates to a novel antigen presentation enhancing hybrid
 CC polypeptide. The novel polypeptide has an N-terminal element consisting
 CC of 4-16 residues of a mammalian II-key peptide and its non-N-terminal
 CC deletion modifications, a chemical structure covalently linking the N-
 CC terminal element to an MHC class II-presented epitope of a C-terminal
 CC element. The C-terminal element comprises an antigenic epitope, which
 CC binds to an antigenic peptide binding site of an MHC class II molecule.

CC The antigen presentation enhancing hybrid polypeptide has the following
 CC activities: antibacterial, virucide, fungicide, antirheumatic,
 CC antihistaminic, neuroprotective, dermatological, immunosuppressive,
 CC antinflammatory, antidiabetic, and antithyroid. The antigen presentation
 CC enhancing hybrid polypeptide is useful for modulating the immune response
 CC in an individual and for treating infections (such as bacteria, virus,
 CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
 CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 XX represents a mammalian II key related peptide epitope of the invention.

SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 8; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 ||||
 Db 7 NGTF 10

RESULT 29
 ADO38565
 ID ADO38565 standard; peptide; 10 AA.
 XX
 AC ADO38565;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Carcinoembryonic antigen (CEA) MHC class II-presented epitope #23.
 XX
 KW Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
 KW Antihistaminic; Neuroprotective; Antinflammatory; Dermatological;
 KW Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;
 KW Antiallergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;
 KW MHC Class II: II-key motif; immune response; anthrax; EBOLA; HIV;
 KW influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
 KW rickettsia; rheumatoid arthritis; multiple sclerosis;
 KW lupus erythematosus; diabetes mellitus; myasthenia gravis;
 KW autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
 KW allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
 KW adenoma; carcinoembryonic antigen; CEA; human.
 XX
 OS Homo sapiens.
 XX
 PN US2004058881-A1.
 XX
 PD 25-MAR-2004.
 XX
 PF 24-SEP-2002; 2002US-00253286.
 XX
 PR 24-SEP-2002; 2002US-00253286.
 XX
 PA (ANTI-) ANTIGEN EXPRESS INC.
 XX
 PI Humphreys RE, Xu M;
 XX
 DR WPI; 2004-294259/27.
 XX
 XX New non-naturally occurring protein or polypeptide modified by
 PT recombinant DNA techniques, useful for treating multiple sclerosis,
 PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
 PT colitis, cancer or psoriasis.
 XX
 PS Example 9; Page 35; 90pp; English.
 XX

CC The invention relates to a non-naturally occurring protein or polypeptide
 CC (I) modified by recombinant DNA techniques comprising: a C-terminal
 CC element comprising an MHC Class II-presented epitope; an N-terminal
 CC element comprising an II-key motif; and an intervening element comprising
 CC a sequence of 4-11 amino acid residues where the modification by
 CC recombinant DNA techniques taking place within elements (b) and (c). Also

described are methods for: suppressing or enhancing an immune response directed toward an MHC (major histocompatibility complex) Class II-presented epitope of interest. Suppressing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding an II-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the II-key motif to decrease its conformance to the archetypal II-key regulatory motif. Enhancing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence lacking an II-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the nucleic acid sequence to introduce an II-key motif appropriately spaced from the MHC Class II-presented epitope. The protein or polypeptide of interest corresponds to a protein or polypeptide encoded by an infectious pathogen selected from anthrax, Ebola, HIV or influenza, preferably vaccinia virus. The non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques is useful for treating infectious diseases caused or associated with infection by a bacterium, virus, parasite, fungus, rickettsia or other infectious agents. It is also useful for treating rheumatoid arthritis, multiple sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis, autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma, allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or adenomas. The present sequence represents the amino acid sequence of carcinoembryonic antigen (CEA) MHC class II-presented epitope used in the invention.

XX SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 7 NGTF 10

RESULT 30
ADP80010
ID ADP80010 standard; peptide; 10 AA.

AC ADP80010;

XX DT 18-NOV-2004 (first entry)

XX Human HLA-A24 epitope vaccine peptide SeqID265.

XX epitope; HLA-A1; HLA-A2; HLA-A3; HLA-A24; HLA-B7; HLA-B44;
XX tumour associated antigen peptide; cytostatic; vaccine; human.

XX Homo sapiens.

XX WO2004052917-A2.

XX 24-JUN-2004.

XX 10-DEC-2003; 2003WO-US038949.

XX 10-DEC-2002; 2002US-0432017P.

XX (EPIM-) EPIMUNE INC.

XX Keogh EA, Southwood S, Fikes JD, Sette A;

XX WPI; 2004-468809/44.

XX New HLA-A1, A2, A3, A24, B7 or B44 tumor associated antigen peptides,
PT useful in preparing a composition for diagnosing or treating tumor
PT associated antigen-related disease.

XX Claim 1; SEQ ID NO 265; 244pp; English.
PS
CC This invention relates to a novel isolated peptide which comprises at least 1, 2, 3, 4, 5, 6 or 7 epitopes or analogues of the epitopes given in the specification. The peptide comprises HLA-A1, A2, A3, A24, B7 or B44 tumour associated antigen peptide. The invention may be useful for the production of compounds with a cytostatic activity or for the production of a vaccine. The peptide is useful in preparing a composition diagnosing or treating tumour associated antigen-related disease. The present sequence is that of an epitope peptide for use in the peptide of the invention.

XX SQ Sequence 10 AA;

Query Match 23.5%; Score 23; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 7 NGTF 10

RESULT 31

AAR67090

ID AAR67090 standard; peptide; 12 AA.

AC AAR67090;

XX DT 25-MAR-2003 (revised)

DT 30-JUN-1995 (first entry)

XX Taste modifier peptide (TM-MIR2) contg. proline brackets.

XX Taste modifier peptide; food; drink; sweet; sour; interaction site;
XX constrained conformation; miraculin; curculin.

OS Synthetic.

XX WO9425482-A1.

XX 10-NOV-1994.

XX 21-APR-1994; 94WO-US004294.

XX 23-APR-1993; 93US-00051741.

XX 29-OCT-1993; 93US-00143364.

XX (EVAN/) EVANS H J.

XX (KINI/) KINI R M.

XX Evans HJ, Kini RM;

XX WPI; 1994-358186/44.

XX Peptide homologue or analogue with constrained conformation - has proline residues flanking the interaction site to impart greater, or more stable, biological activity.

XX Example 2; Page 33; 57pp; English.

XX AAR67089-91 are sweet peptides derived from naturally occurring polypeptides that contain proline or proline/cysteine brackets. These peptides are shortened to form fragments that contain one or more interaction sites of interest. AAR67089-90 are deriv. from miraculin. The taste modifier peptides probably work through interaction with taste receptors, modifying sour taste into sweet taste. Structural constraints, partic. cyclisation of the peptides, may help in the heat stability of these peptides. Stabilisation should increase the usefulness of these peptides in cooking. The data collected demonstrates that interaction sites possess activity when present in a polypeptide that differs from the native form. Inclusion of conformation-constraining moieties can have

CC desirable effects on an interaction site. (Also see AAR67011-88 and
 CC AAR67092-152 for analogues of other biologically active peptides contg.
 CC an interaction site flanked by conformation constraining gps., eg. RGD
 CC peptides.) (updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 12 AA;
 Query Match 23.5%; Score 23; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 33 NGTF 36
 DB 4 NGTF 7
 RESULT 32
 ID ABU11918 standard; peptide; 12 AA.
 XX
 AC ABU11918;
 DT 13-FEB-2003 (first entry)
 XX Human HGPRBW11 asparagine glycosylation site #1.
 DE
 XX Human; immunogen; HGPRBW11; HGPRBW11v1; HGPRBW11v2; GPCR74; GPCR81;
 KW G-protein coupled receptor; cardiovascular disease; arrhythmia;
 KW myocardial infarction; congestive heart failure; cardiomyopathy;
 KW atherosclerosis; arteriosclerosis; embolism; angina; thrombosis;
 KW hypertension; Alzheimer's disease; Parkinson's disease; osteoporosis;
 KW obesity; human immunodeficiency virus infection; HIV; schizophrenia;
 KW sleeplessness; acquired immunodeficiency syndrome; AIDS; leukaemia;
 KW sepsis; inflammation; psoriasis; Gaucher's disease; ischaemia.
 XX Homo sapiens.
 OS
 XX
 XX W0200286123-A2.
 XX 31-OCT-2002.
 XX 16-NOV-2001; 2001WO-US044019.
 XX 17-NOV-2000; 2000US-0249613P.
 XX 21-DEC-2000; 2000US-0257611P.
 XX 16-JUL-2001; 2001US-0305818P.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Feder J, Nelson TC, Ramanathan C, Cacace AM, Barber LE;
 PI WPI; 2003-093137/08.
 DR
 XX New human G-protein coupled receptor HGPRBW11 polypeptide or
 PT polynucleotide, useful for preventing, treating or ameliorating e.g.
 PT myocardial infarction, angina, thrombosis, Alzheimer's disease,
 PT schizophrenia, AIDS, leukemia.
 XX
 XX Disclosure; Page 51; 444pp; English.
 PS
 XX The invention relates to an isolated polypeptide (designated HGPRBW11),
 CC which has a G-protein coupled receptor (GPCR) activity (also known as
 CC GPCR74 or GPCR81) and is encoded by the cDNA contained in ARCC Deposit
 CC Number PTA-2766, its variants (HGPRBW11v1 and HGPRBW11v2), fragments,
 CC domains, species homologues and proteins 95% similar to it. Also included
 CC are: (1) the nucleic acids encoding the HGPRBW11 proteins (including
 CC variants, fragments, complements and sequences 95% similar to the
 CC HGPRBW11 nucleic acids); (2) recombinant vectors; (3) transformed host
 CC cells expressing HGPRBW11; (4) an anti-HGPRBW11 antibody; (5)
 CC diagnosing a pathological condition or a susceptibility to a pathological
 CC condition in a subject; (6) identifying a binding partner to the
 CC HGPRBW11 polypeptide; (7) identifying an activity in a biological assay;
 CC (8) a process for making polynucleotide sequences encoding a gene product

CC having altered GPCR activity; (9) a shuffled polynucleotide produced by
 CC the method of (8); and (10) screening for candidate compounds capable of
 CC binding to and/or modulating activity of a GPCR. The polypeptide or the
 CC polynucleotide is useful for preventing, treating or ameliorating a
 CC medical condition, particularly cardiovascular diseases or disorders,
 CC e.g. myocardial infarction, congestive heart failure, arrhythmias,
 CC cardiomyopathy, atherosclerosis, arteriosclerosis, embolism, angina,
 CC thrombosis or hypertension). The HGPRBW11 polypeptide or polynucleotide
 CC is also useful for treating, preventing or ameliorating e.g. Alzheimer's
 CC disease, Parkinson's disease, osteoporosis, obesity, human
 CC immunodeficiency virus (HIV) infections, schizophrenia, sleeplessness,
 CC acquired immunodeficiency syndrome (AIDS), leukaemia, sepsis,
 CC inflammations, psoriasis, Gaucher's disease or ischaemia (many other
 CC diseases and disorders are listed in the specification). The present
 CC sequence represents a protein motif or domain of an HGPRBW11 protein (or
 CC variant) which may be used as an immunogen
 XX
 SQ Sequence 12 AA;
 Query Match 23.5%; Score 23; DB 6; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 33 NGTF 36
 DB 4 NGTF 7
 RESULT 33
 ID AAB88073 standard; peptide; 14 AA.
 XX
 AC AAB88073;
 DT 17-MAY-2001 (first entry)
 XX CD66 peptide CD66c-18.
 DE
 XX CD66; CEACAM; adhesion molecule; antiviral; antibacterial;
 KW antiinflammatory; cytostatic; neutrophil activation; proliferation;
 KW differentiation; cancer; angiogenesis.
 XX Unidentified.
 OS
 XX W0200113937-A1.
 XX 01-MAR-2001.
 XX 25-AUG-2000; 2000WO-US023482.
 XX 26-AUG-1999; 99US-0150791P.
 XX 02-SEP-1999; 99US-0152501P.
 XX (SKUB/) SKUBITZ K M.
 PA (SKUB/) SKUBITZ A P N.
 XX Skubitz KM, Skubitz APN;
 PI WPI; 2001-234981/24.
 DR
 XX Novel peptides useful for activating neutrophils or blocking activation
 PT of neutrophils, modulating homotypic or heterotypic adhesion of CD66
 PT polypeptides, and modulating immune cell activation.
 XX
 PS Claim 1; Page 45; 102pp; English.
 XX The present sequence is an isolated peptide that was tested for its
 CC ability to modulate the function of CD66 family polypeptides and CD66
 CC ligands. 106 sequences of 13 or 14 amino acids in length, and their
 CC analogues, were identified that modulate the function of at least one
 CC CD66 family polypeptide and/or at least one ligand of the polypeptide.
 CC The peptides are capable of modulating activation of neutrophils,
 CC activation or inhibition, proliferation and/or differentiation of T-

CC cells, B-cells, NK cells, LAK cells, dendritic cells, or other immune
 CC system cells, proliferation and/or differentiation of epithelial cells,
 CC homotypic and/or heterotypic adhesion among CD66 family polypeptides and
 CC adhesion of CD66 family polypeptides to other ligands. The peptides are
 CC useful for delivering a therapeutically active agent to a patient, for
 CC modifying the metastasis of malignant cells, for altering bacterial or
 CC viral binding to cells or a biomaterial, for altering cell adhesion to a
 CC biomaterial, for detecting tumours, for detecting inflammation, for
 CC detecting a CD66 protein or its ligand, for altering angiogenesis by
 CC contacting endothelial cells, tumour cells or immune cells, for altering
 CC an immune response, and for altering keratinocyte proliferation
 XX Sequence 14 AA;
 SQ

Query Match 23.5%; Score 23; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 Db |||||
 3 NGTF 6

RESULT 34
 AAB88018
 ID AAB88018 standard; peptide; 14 AA.
 AC AAB88018;
 XX
 XX 17-MAY-2001 (first entry)
 DT
 XX CD66 peptide CD66a-18.
 DE
 XX CD66; CEACAM; adhesion molecule; antiviral; antibacterial;
 KW antiinflammatory; cytostatic; neutrophil activation; proliferation;
 KW differentiation; cancer; angiogenesis.
 XX Unidentified.
 OS
 XX WO200113937-A1.
 PN
 XX 01-MAR-2001.
 PD
 XX 25-AUG-2000; 2000WO-US023482.
 PF
 XX 26-AUG-1999; 99US-0150791P.
 PR
 XX 02-SEP-1999; 99US-0152501P.
 PR
 XX (SKUB/) SKUBITZ K M.
 PA (SKUB/) SKUBITZ A P N.
 PA Skubitz KM, Skubitz APN;
 PI WPI; 2001-234981/24.
 DR
 XX Novel peptides useful for activating neutrophils or blocking activation
 PT of neutrophils, modulating homotypic or heterotypic adhesion of CD66
 PT polypeptides, and modulating immune cell activation.
 XX Claim 1; Page 39; 102pp; English.

XX The present sequence is an isolated peptide that was tested for its
 CC ability to modulate the function of CD66 family polypeptides and CD66
 CC ligands. 106 sequences of 13 or 14 amino acids in length, and their
 CC analogues, were identified that modulate the function of at least one
 CC CD66 family polypeptide and/or at least one ligand of the polypeptide.
 CC The peptides are capable of modulating activation of neutrophils,
 CC activation or inhibition, proliferation and/or differentiation of T-
 CC cells, B-cells, NK cells, LAK cells, dendritic cells, or other immune
 CC system cells, proliferation and/or differentiation of epithelial cells,
 CC homotypic and/or heterotypic adhesion among CD66 family polypeptides and
 CC adhesion of CD66 family polypeptides to other ligands. The peptides are
 CC useful for delivering a therapeutically active agent to a patient, for
 CC modifying the metastasis of malignant cells, for altering bacterial or
 CC viral binding to cells or a biomaterial, for altering cell adhesion to a
 CC biomaterial, for detecting tumours, for detecting inflammation, for
 CC detecting a CD66 protein or its ligand, for altering angiogenesis by
 CC contacting endothelial cells, tumour cells or immune cells, for altering

CC modifying the metastasis of malignant cells, for altering bacterial or
 CC viral binding to cells or a biomaterial, for altering cell adhesion to a
 CC biomaterial, for detecting tumours, for detecting inflammation, for
 CC detecting a CD66 protein or its ligand, for altering angiogenesis by
 CC contacting endothelial cells, tumour cells or immune cells, for altering
 CC an immune response, and for altering keratinocyte proliferation
 XX Sequence 14 AA;
 SQ

Query Match 23.5%; Score 23; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 Db |||||
 3 NGTF 6

RESULT 35
 AAB88099
 ID AAB88099 standard; peptide; 14 AA.
 AC AAB88099;
 XX
 XX 17-MAY-2001 (first entry)
 DT
 XX CD66 peptide CD66e-20.
 DE
 XX CD66; CEACAM; adhesion molecule; antiviral; antibacterial;
 KW antiinflammatory; cytostatic; neutrophil activation; proliferation;
 KW differentiation; cancer; angiogenesis.
 XX Unidentified.
 OS
 XX WO200113937-A1.
 PN
 XX 01-MAR-2001.
 PD
 XX 25-AUG-2000; 2000WO-US023482.
 PF
 XX 26-AUG-1999; 99US-0150791P.
 PR
 XX 02-SEP-1999; 99US-0152501P.
 PR
 XX (SKUB/) SKUBITZ K M.
 PA (SKUB/) SKUBITZ A P N.
 PA Skubitz KM, Skubitz APN;
 PI WPI; 2001-234981/24.
 DR
 XX Novel peptides useful for activating neutrophils or blocking activation
 PT of neutrophils, modulating homotypic or heterotypic adhesion of CD66
 PT polypeptides, and modulating immune cell activation.
 XX Claim 1; Page 48; 102pp; English.

XX The present sequence is an isolated peptide that was tested for its
 CC ability to modulate the function of CD66 family polypeptides and CD66
 CC ligands. 106 sequences of 13 or 14 amino acids in length, and their
 CC analogues, were identified that modulate the function of at least one
 CC CD66 family polypeptide and/or at least one ligand of the polypeptide.
 CC The peptides are capable of modulating activation of neutrophils,
 CC activation or inhibition, proliferation and/or differentiation of T-
 CC cells, B-cells, NK cells, LAK cells, dendritic cells, or other immune
 CC system cells, proliferation and/or differentiation of epithelial cells,
 CC homotypic and/or heterotypic adhesion among CD66 family polypeptides and
 CC adhesion of CD66 family polypeptides to other ligands. The peptides are
 CC useful for delivering a therapeutically active agent to a patient, for
 CC modifying the metastasis of malignant cells, for altering bacterial or
 CC viral binding to cells or a biomaterial, for altering cell adhesion to a
 CC biomaterial, for detecting tumours, for detecting inflammation, for
 CC detecting a CD66 protein or its ligand, for altering angiogenesis by
 CC contacting endothelial cells, tumour cells or immune cells, for altering

CC an immune response, and for altering keratinocyte proliferation
 XX
 SQ Sequence 14 AA;

Query Match 23.5%; Score 23; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 3 NGTF 6

RESULT 36
 AAB88056
 ID AAB88056 standard; peptide; 14 AA.

XX AC AAB88056;
 XX DT 17-MAY-2001 (first entry)
 XX DE CD66 peptide CD66b-18.

XX CD66; CEACAM; adhesion molecule; antiviral; antibacterial;
 XX antiinflammatory; cytostatic; neutrophil activation; proliferation;
 XX differentiation; cancer; angiogenesis.

XX OS Unidentified.

XX WO200113937-A1.

XX PD 01-MAR-2001.

XX PF 25-AUG-2000; 2000WO-US023482.

XX PR 26-AUG-1999; 99US-0150791P.

XX PR 02-SEP-1999; 99US-0152501P.

XX PA (SKUB/) SKUBITZ K M.

XX PA (SKUB/) SKUBITZ A P N.

XX PI Skubitz KM, Skubitz APN;

XX WPI; 2001-234981/24.

XX Novel peptides useful for activating neutrophils or blocking activation
 XX of neutrophils, modulating homotypic or heterotypic adhesion of CD66
 XX polypeptides, and modulating immune cell activation.

XX Claim 1; Page 43; 102pp; English.

XX The present sequence is an isolated peptide that was tested for its
 CC ability to modulate the function of CD66 family polypeptides and CD66
 CC ligands. 106 sequences of 13 or 14 amino acids in length, and their
 CC analogues, were identified that modulate the function of at least one
 CC CD66 family polypeptide and/or at least one ligand of the polypeptide.
 CC The peptides are capable of modulating activation of neutrophils,
 CC activation or inhibition, proliferation and/or differentiation of T-
 CC cells, B-cells, NK cells, LAK cells, dendritic cells, or other immune
 CC system cells, proliferation and/or differentiation of epithelial cells,
 CC homotypic and/or heterotypic adhesion among CD66 family polypeptides and
 CC adhesion of CD66 family polypeptides to other ligands. The peptides are
 CC useful for delivering a therapeutically active agent to a patient, for
 CC modifying the metastasis of malignant cells, for altering bacterial or
 CC viral binding to cells or a biomaterial, for altering cell adhesion to a
 CC biomaterial, for detecting tumours, for detecting inflammation, for
 CC detecting a CD66 protein or its ligand, for altering angiogenesis by
 CC contacting endothelial cells, tumour cells or immune cells, for altering
 CC an immune response, and for altering keratinocyte proliferation

XX Sequence 14 AA;

Query Match 23.5%; Score 23; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 3 NGTF 6

RESULT 37

ABP67248
 ID ABP67248 standard; peptide; 14 AA.

XX AC ABP67248;

XX DT 10-DEC-2002 (first entry)

XX DE Human CD66 family modulating peptide SEQ ID NO 119.

XX Human; CD66; CEACAM; cytostatic; antiinflammatory; immunomodulator;
 XX antibacterial; virucide; gene therapy; vaccine; neutrophil;
 XX immune system; autoimmune disease; cancer; infection; bacterial; virus;
 XX inflammatory disease; transplantation; immunisation.

XX OS Homo sapiens.

XX PN WO200268601-A2.

XX PD 06-SEP-2002.

XX PF 27-FEB-2002; 2002WO-US005720.

XX PR 28-FEB-2001; 2001US-0272113P.

XX PA (SKUB/) SKUBITZ K M.

XX PA (SKUB/) SKUBITZ A P N.

XX PI Skubitz KM, Skubitz APN;

XX WPI; 2002-706981/76.

XX New peptide from a surface exposed region of a CD66 family member, useful
 XX for modulating the function of CD66 family members, e.g. activation of
 XX neutrophils, for treating or diagnosing autoimmune diseases or cancer,
 XX and as a vaccine.

XX Disclosure; Page 15; 96pp; English.

XX The invention relates to an isolated peptide (I) from a surface exposed
 XX region of a CD66 family member (ABP67130-ABP67990). The peptide
 XX modulates: (a) activation of neutrophils; (b) activation or inhibition of
 XX T-cells, B-cells, NK cells, LAK cells, dendritic cells or other immune
 XX system cells; (c) proliferation and/or differentiation of the above
 XX cells, including epithelial cells; (d) homotypic and/or heterotypic
 XX adhesion among CD66 family members; and (e) adhesion of CD66 family
 XX members to other ligands. The peptide is useful in modulating the
 XX function of CD66 family members and/or their ligands, such as activation
 XX of neutrophils and activation, inhibition, proliferation and/or
 XX differentiation of the immune cells. The peptides may also be used in
 XX treating or diagnosing autoimmune diseases, cancer, infections (e.g.
 XX bacterial or viral) or inflammatory diseases, in transplantation
 XX therapies and for immunisation

XX Sequence 14 AA;

Query Match 23.5%; Score 23; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 ||||
 Db 3 NGTF 6

RESULT 38
 ABUI1935
 ID ABUI1935 standard; peptide; 14 AA.
 AC ABUI1935;
 XX
 XX 13-FEB-2003 (first entry)
 DT
 XX Human HGPRBMY11v1 asparagine glycosylation site #1.
 DE
 XX
 XX Human; immungen; HGPRBMY11; HGPRBMY11v1; HGPRBMY11v2; GPCR74; GPCR81;
 KW G-protein coupled receptor; cardiovascular disease; arrhythmia;
 KW myocardial infarction; congestive heart failure; cardiomyopathy;
 KW atherosclerosis; arteriosclerosis; embolism; angina; thrombosis;
 KW hypertension; Alzheimer's disease; Parkinson's disease; osteoporosis;
 KW obesity; human immunodeficiency virus infection; HIV; schizophrenia;
 KW sleeplessness; acquired immunodeficiency syndrome; AIDS; leukaemia;
 KW sepsis; inflammation; psoriasis; Gaucher's disease; ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200286123-A2.
 PN
 XX 31-OCT-2002.
 PD
 XX 16-NOV-2001; 2001WO-US044019.
 PF
 XX 17-NOV-2000; 2000US-0249613P.
 PR
 XX 21-DEC-2000; 2000US-0257611P.
 PR
 XX 16-JUL-2001; 2001US-0305818P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Feder J, Nelson TC, Ramanathan C, Cacace AM, Barber LE;
 PI
 XX WPI; 2003-093137/08.
 DR
 XX New human G-protein coupled receptor HGPRBMY11 polypeptide or
 PT polynucleotide, useful for preventing, treating or ameliorating e.g.
 PT myocardial infarction, angina, thrombosis, Alzheimer's disease,
 PT schizophrenia, AIDS, leukemia.
 XX
 PS Disclosure; Page 79; 444pp; English.
 CC The invention relates to an isolated polypeptide (designated HGPRBMY11),
 CC which has a G-protein coupled receptor (GPCR) activity (also known as
 CC GPCR74 or GPCR81) and is encoded by the cDNA contained in ATCC Deposit
 CC Number PTA-2766, its variants (HGPRBMY11v1 and HGPRBMY11v2), fragments,
 CC domains, species homologues and proteins 95% similar to it. Also included
 CC are: (1) the nucleic acids encoding the HGPRBMY11 proteins (including
 CC variants, fragments, complements and sequences 95% similar to the
 CC HGPRBMY11 nucleic acids); (2) recombinant vectors; (3) transformed host
 CC cells expressing HGPRBMY11; (4) an anti-HGPRBMY11 antibody; (5)
 CC diagnosing a pathological condition or a susceptibility to a pathological
 CC condition in a subject; (6) identifying a binding partner to the
 CC HGPRBMY11 polypeptide; (7) identifying an activity in a biological assay;
 CC (8) a process for making polynucleotide sequences encoding a gene product
 CC having altered GPCR activity; (9) a shuffled polynucleotide produced by
 CC binding to and/or modulating activity of a GPCR. The polypeptide or the
 CC polynucleotide is useful for preventing, treating or ameliorating a
 CC medical condition, particularly cardiovascular diseases or disorders,
 CC e.g. myocardial infarction, congestive heart failure, arrhythmias,
 CC cardiomyopathy, atherosclerosis, arteriosclerosis, embolism, angina,
 CC thrombosis or hypertension). The HGPRBMY11 polypeptide or polynucleotide
 CC is also useful for treating, preventing or ameliorating e.g. Alzheimer's
 CC disease, Parkinson's disease, osteoporosis, obesity, human
 CC immunodeficiency virus (HIV) infections, schizophrenia, sleeplessness,
 CC acquired immunodeficiency syndrome (AIDS), leukaemia, sepsis,
 CC inflammations, psoriasis, Gaucher's disease or ischaemia (many other
 CC diseases and disorders are listed in the specification). The present
 CC sequence represents a protein motif or domain of an HGPRBMY11 protein (or
 CC variant) which may be used as an immungen

XX
 SQ Sequence 14 AA;
 Query Match 23.5%; Score 23; DB 6; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 NGTF 36
 Db
 6 NGTF 9
 RESULT 39
 ABUI1953
 ID ABUI1953 standard; peptide; 14 AA.
 XX
 AC ABUI1953;
 XX
 XX 13-FEB-2003 (first entry)
 DT
 XX Human HGPRBMY11v1 casein kinase II phosphorylation site #2.
 DE
 XX
 XX Human; immungen; HGPRBMY11; HGPRBMY11v1; HGPRBMY11v2; GPCR74; GPCR81;
 KW G-protein coupled receptor; cardiovascular disease; arrhythmia;
 KW myocardial infarction; congestive heart failure; cardiomyopathy;
 KW atherosclerosis; arteriosclerosis; embolism; angina; thrombosis;
 KW hypertension; Alzheimer's disease; Parkinson's disease; osteoporosis;
 KW obesity; human immunodeficiency virus infection; HIV; schizophrenia;
 KW sleeplessness; acquired immunodeficiency syndrome; AIDS; leukaemia;
 KW sepsis; inflammation; psoriasis; Gaucher's disease; ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200286123-A2.
 PN
 XX 31-OCT-2002.
 PD
 XX 16-NOV-2001; 2001WO-US044019.
 PF
 XX 17-NOV-2000; 2000US-0249613P.
 PR
 XX 21-DEC-2000; 2000US-0257611P.
 PR
 XX 16-JUL-2001; 2001US-0305818P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Feder J, Nelson TC, Ramanathan C, Cacace AM, Barber LE;
 PI
 XX WPI; 2003-093137/08.
 DR
 XX New human G-protein coupled receptor HGPRBMY11 polypeptide or
 PT polynucleotide, useful for preventing, treating or ameliorating e.g.
 PT myocardial infarction, angina, thrombosis, Alzheimer's disease,
 PT schizophrenia, AIDS, leukemia.
 XX
 PS Disclosure; Page 78; 444pp; English.
 CC The invention relates to an isolated polypeptide (designated HGPRBMY11),
 CC which has a G-protein coupled receptor (GPCR) activity (also known as
 CC GPCR74 or GPCR81) and is encoded by the cDNA contained in ATCC Deposit
 CC Number PTA-2766, its variants (HGPRBMY11v1 and HGPRBMY11v2), fragments,
 CC domains, species homologues and proteins 95% similar to it. Also included
 CC are: (1) the nucleic acids encoding the HGPRBMY11 proteins (including
 CC variants, fragments, complements and sequences 95% similar to the
 CC HGPRBMY11 nucleic acids); (2) recombinant vectors; (3) transformed host
 CC cells expressing HGPRBMY11; (4) an anti-HGPRBMY11 antibody; (5)
 CC diagnosing a pathological condition or a susceptibility to a pathological
 CC condition in a subject; (6) identifying a binding partner to the
 CC HGPRBMY11 polypeptide; (7) identifying an activity in a biological assay;
 CC (8) a process for making polynucleotide sequences encoding a gene product
 CC having altered GPCR activity; (9) a shuffled polynucleotide produced by
 CC binding to and/or modulating activity of a GPCR. The polypeptide or the
 CC polynucleotide is useful for preventing, treating or ameliorating a
 CC medical condition, particularly cardiovascular diseases or disorders,
 CC e.g. myocardial infarction, congestive heart failure, arrhythmias,
 CC cardiomyopathy, atherosclerosis, arteriosclerosis, embolism, angina,
 CC thrombosis or hypertension). The HGPRBMY11 polypeptide or polynucleotide
 CC is also useful for treating, preventing or ameliorating e.g. Alzheimer's
 CC disease, Parkinson's disease, osteoporosis, obesity, human
 CC immunodeficiency virus (HIV) infections, schizophrenia, sleeplessness,
 CC acquired immunodeficiency syndrome (AIDS), leukaemia, sepsis,
 CC inflammations, psoriasis, Gaucher's disease or ischaemia (many other
 CC diseases and disorders are listed in the specification). The present
 CC sequence represents a protein motif or domain of an HGPRBMY11 protein (or
 CC variant) which may be used as an immungen

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:14:52 ; Search time 28.2276 Seconds
(without alignments)
148.094 Million cell updates/sec

Title: US-09-936-956-1
Perfect score: 98
Sequence: 1 XGXXXXXXXXXXXXXAX.....XXXXXXXXXXXXXG 56

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	27.6	459	4	US-09-902-540-10583
2	25	25.5	103	4	US-09-902-540-13360
3	25	25.5	1026	1	US-08-194-290-7
4	25	25.5	1026	2	US-08-614-377A-7
5	25	25.5	1026	3	US-09-142-648B-7
6	24	24.5	418	2	US-08-873-479-44
7	23	23.5	12	2	US-08-934-222-83
8	23	23.5	12	2	US-08-933-402-83
9	23	23.5	12	2	US-09-207-621-83
10	23	23.5	12	2	US-08-532-818-83
11	23	23.5	12	3	US-09-231-797-83
12	23	23.5	12	3	US-08-934-224-83
13	23	23.5	12	3	US-08-933-843-83
14	23	23.5	12	3	US-08-934-223-83
15	23	23.5	12	3	US-09-413-432-83
16	23	23.5	23	1	US-08-258-639A-9
17	23	23.5	23	2	US-08-900-951-9
18	23	23.5	23	4	US-09-270-767-60291
19	23	23.5	23	5	PCT-US95-07391A-9
20	23	23.5	31	3	US-09-209-799D-22
21	23	23.5	31	4	US-09-997-792A-19
22	23	23.5	38	3	US-09-082-279B-1420
23	23	23.5	38	3	US-09-315-304B-1420
24	23	23.5	38	4	US-09-834-784-1420
25	23	23.5	38	4	US-09-515-965A-1420
26	23	23.5	38	4	US-09-350-641C-1420
27	23	23.5	38	4	US-09-350-841A-1420

28	23	23.5	39	3	US-09-082-279B-1419	Sequence 1419, Ap
29	23	23.5	39	3	US-09-315-304B-1419	Sequence 1419, Ap
30	23	23.5	39	4	US-09-834-784-1419	Sequence 1419, Ap
31	23	23.5	39	4	US-09-515-965A-1419	Sequence 1419, Ap
32	23	23.5	39	4	US-09-350-641C-1419	Sequence 1419, Ap
33	23	23.5	39	4	US-09-350-841A-1419	Sequence 1419, Ap
34	23	23.5	40	3	US-09-082-279B-1242	Sequence 1242, Ap
35	23	23.5	40	3	US-09-315-304B-1242	Sequence 1242, Ap
36	23	23.5	40	4	US-09-834-784-1242	Sequence 1242, Ap
37	23	23.5	40	4	US-09-515-965A-1242	Sequence 1242, Ap
38	23	23.5	40	4	US-09-515-965A-1985	Sequence 1985, Ap
39	23	23.5	40	4	US-09-350-641C-1242	Sequence 1242, Ap
40	23	23.5	40	4	US-09-350-641C-1747	Sequence 1747, Ap
41	23	23.5	40	4	US-09-350-641C-1748	Sequence 1748, Ap
42	23	23.5	40	4	US-09-350-841A-1242	Sequence 1242, Ap
43	23	23.5	40	4	US-09-350-841A-1936	Sequence 1936, Ap
44	23	23.5	40	4	US-09-350-841A-1937	Sequence 1937, Ap
45	23	23.5	40	4	US-09-350-841A-1937	Sequence 1937, Ap

ALIGNMENTS

RESULT 1

US-09-902-540-10583
; Sequence 10583, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10583
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10583

Query Match 27.6%; Score 27; DB 4; Length 459;
Best Local Similarity 23.2%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXXXXXXXXXXG 56
Db 87 NGTFDGRQVVQVAYLTSGDKVELG 110

RESULT 2

US-09-902-540-13360
; Sequence 13360, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13360
; LENGTH: 103
; TYPE: PRT

```
; ORGANISM: Myxococcus xanthus
US-09-902-540-13360

Query Match      25.5%; Score 25; DB 4; Length 103;
Best Local Similarity 29.2%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXGDXGDXG 56
    |||||
Db 69 NGTFVNGEKVKQSLKRGDRILIG 92

RESULT 3
US-08-194-290-7
; Sequence 7, Application US/08194290
; Patent No. 5500353
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H
; TITLE OF INVENTION: Bacterial surface protein expression
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shlesinger, Arkwright & Garvey
; STREET: 3000 South Eads Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,290
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garvey, George A
; REGISTRATION NUMBER: 17737
; REFERENCE/DOCKET NUMBER: 5946-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-5600
; TELEFAX: 703-836-5288
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-290-7

Query Match      25.5%; Score 25; DB 1; Length 1026;
Best Local Similarity 25.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXGDXGDXG 56
    |||||
Db 267 NDTFVAGEVAGATLTVGDTLSGG 290

RESULT 4
US-08-614-377A-7
; Sequence 7, Application US/08614377A
; Patent No. 5978664
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H
; APPLICANT: No. 5976864ellini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF
; TITLE OF INVENTION: HETEROLOGOUS
; TITLE OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson PC
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/614,377A
; FILING DATE: 12-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/194,290
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/895,367
; FILING DATE: 09-JUNE-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsao, Y. Rocky
; REGISTRATION NUMBER: 34053
; REFERENCE/DOCKET NUMBER: 08106/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-614-377A-7

Query Match      25.5%; Score 25; DB 2; Length 1026;
Best Local Similarity 25.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXGDXGDXG 56
    |||||
Db 267 NDTFVAGEVAGATLTVGDTLSGG 290

RESULT 5
US-09-142-648B-7
; Sequence 7, Application US/09142648B
; Patent No. 6210948
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H
; APPLICANT: No. 6210948ellini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CAULO
; FILE REFERENCE: 08106/002002
; CURRENT APPLICATION NUMBER: US/09/142,648B
; CURRENT FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: PCT/CA97/00167
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: US 07/614,377
; PRIOR FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-09-142-648B-7

Query Match      25.5%; Score 25; DB 3; Length 1026;
```

Best Local Similarity 25.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
Db 267 NDTFVAGEVAGAAATLTGVDTLSSG 290

RESULT 6

US-08-873-479-44
; Sequence 44, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 58917010th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agria, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-44

Query Match 24.5%; Score 24; DB 2; Length 418;
Best Local Similarity 26.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDX 51
Db 307 NGTYRVADFSRGGHKTAGD 325

RESULT 7

US-08-934-222-83
; Sequence 83, Application US/08934222
; Patent No. 5928896
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,222
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-934-222-83

Query Match 23.5%; Score 23; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 4 NGTF 7

RESULT 8

US-08-933-402-83
; Sequence 83, Application US/08933402
; Patent No. 5948887
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,402
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-933-402-83

Query Match 23.5%; Score 23; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db ||||
4 NGTF 7

RESULT 9

US-09-207-621-83
; Sequence 83, Application US/09207621
; Patent No. 5952465
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/207,621
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-207-621-83

Query Match 23.5%; Score 23; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-207-621-83

Query Match 23.5%; Score 23; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db ||||
4 NGTF 7

RESULT 10

US-08-532-818-83
; Sequence 83, Application US/08532818
; Patent No. 5965698
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,818
; FILING DATE: 03-MAY-1996
; PRIOR APPLICATION DATA: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-532-818-83

Query Match 23.5%; Score 23; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db ||||
4 NGTF 7

RESULT 11

US-09-231-797-83
; Sequence 83, Application US/09231797
; Patent No. 6084066
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner

STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,797
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-231-797-83

Query Match 23.5%; Score 23; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 4 NGTF 7

RESULT 12
US-08-934-224-83
Sequence 83, Application US/08934224
Patent No. 610044
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,224
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996

APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-934-224-83

Query Match 23.5%; Score 23; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 4 NGTF 7

RESULT 13
US-08-933-843-83
Sequence 83, Application US/08933843
Patent No. 611069
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,843
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-933-843-83

Query Match 23.5%; Score 23; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 4 NGTF 7

RESULT 14
US-08-934-223-83
; Sequence 83, Application US/08934223
; Patent No. 6147189
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.
; APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; Constraining Groups Which Flank A Protein-Protein Interaction
; TITLE OF INVENTION: Site
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,223
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-934-223-83

Query Match 23.5%; Score 23; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 4 NGTF 7

RESULT 15
US-09-413-492-83
; Sequence 83, Application US/09413492
; Patent No. 6258550
; GENERAL INFORMATION:
; APPLICANT: EVANS, Herbert J.

APPLICANT: KINI, R. Manjunatha
; TITLE OF INVENTION: Polypeptides That Include Conformation-
; Constraining Groups Which Flank A Protein-Protein Interaction
; TITLE OF INVENTION: Site
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/413,492
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/532,818
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: PCT/US94/04294
; FILING DATE: 21-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/143,364
; FILING DATE: 29-OCT-1993
; APPLICATION NUMBER: U.S. 08/051,741
; FILING DATE: 23-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Isaacson, John P.
; REGISTRATION NUMBER: 33,751
; REFERENCE/DOCKET NUMBER: 040433/0148
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-413-492-83

Query Match 23.5%; Score 23; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 4 NGTF 7

RESULT 16
US-08-258-639A-9
; Sequence 9, Application US/08258639A
; Patent No. 5681733
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/258,639A
 ; FILING DATE: 10 JUNE 1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Healey, William J.
 ; REGISTRATION NUMBER: 36,160
 ; REFERENCE/DOCKET NUMBER: 104385.116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)942-8400
 ; TELEFAX: (202)942-8484
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-258-639A-9

Query Match 23.5%; Score 23; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 DB 16 NGTF 19

RESULT 17

US-08-900-951-9
 ; Sequence 9, Application US/08900951
 ; Patent No. 5919693

GENERAL INFORMATION:

; APPLICANT: Su, Hongsheng
 ; APPLICANT: Blain, Francoise
 ; APPLICANT: Bennett, Clark
 ; APPLICANT: Gu, Kangfu
 ; APPLICANT: Zimmermann, Joseph
 ; APPLICANT: Musil, Roy
 ; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
 ; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
 ; TITLE OF INVENTION: Flavobacterium heparinum
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hale and Dorr
 ; STREET: 1455 Pennsylvania Avenue, N.W.
 ; CITY: Washington, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004

; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC DOS/MS DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/900,951
 ; FILING DATE:
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/258,639
 ; FILING DATE: 10 JUNE 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Healey, William J.
 ; REGISTRATION NUMBER: 36,160
 ; REFERENCE/DOCKET NUMBER: 104385.116
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)942 8400
 ; TELEFAX: (202)942 8484
 ; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-900-951-9

Query Match 23.5%; Score 23; DB 2; Length 23;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 DB 16 NGTF 19

RESULT 18

US-09-270-767-60291
 ; Sequence 60291, Application US/09270767
 ; Patent No. 6703491

GENERAL INFORMATION:

; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 60291
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 ; US-09-270-767-60291

Query Match 23.5%; Score 23; DB 4; Length 23;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
 DB 14 NGTF 17

RESULT 19

PCT-US95-07391A-9

; Sequence 9, Application PC/TUS9507391A

GENERAL INFORMATION:

; APPLICANT: IBEX TECHNOLOGIES and
 ; APPLICANT: ZIMMERMANN, Joseph
 ; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
 ; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
 ; TITLE OF INVENTION: Flavobacterium heparinum
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hale and Dorr
 ; STREET: 1455 Pennsylvania Avenue, N.W.
 ; CITY: Washington, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/07391A
 ; FILING DATE: 09-JUNE-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/258,639
 ; FILING DATE: 10 JUNE 1994

```
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 104385.116PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8400
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-07391A-9

Query Match 23.5%; Score 23; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 16 NGTF 19

RESULT 20
US-09-209-799D-22
; Sequence 22, Application US/09209799D
; Patent No. 6380357
; GENERAL INFORMATION:
; APPLICANT: Hermeling, Ronald
; APPLICANT: Hoffmann, James
; APPLICANT: Narasimhan, Chakravarthy
; TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
; FILE REFERENCE: X-10242
; CURRENT APPLICATION NUMBER: US/09/209,799D
; CURRENT FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-209-799D-22

Query Match 23.5%; Score 23; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 3 NGTF 6

RESULT 21
US-09-997-792A-19
; Sequence 19, Application US/09997792A
; Patent No. 655521
; GENERAL INFORMATION:
; APPLICANT: ELI LILLY and COMPANY
; TITLE OF INVENTION: Glucagon-Like Peptide-1 Crystals
; FILE REFERENCE: X-10242A
; CURRENT APPLICATION NUMBER: US/09/997,792A
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/069,728
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence

; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 104385.116PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8400
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-07391A-9

Query Match 23.5%; Score 23; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 3 NGTF 6

RESULT 22
US-09-082-279B-1420
; Sequence 1420, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1420

Query Match 23.5%; Score 23; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 12 NGTF 15

RESULT 23
US-09-315-304B-1420
; Sequence 1420, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1420
```



```
Query Match      23.5%; Score 23; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      12 NGTF 15

RESULT 24
US-09-834-784-1420
; Sequence 1420, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-834-784-1420

Query Match      23.5%; Score 23; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      12 NGTF 15

RESULT 25
US-09-515-965A-1420
; Sequence 1420, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515,965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide

Query Match      23.5%; Score 23; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      12 NGTF 15

RESULT 26
US-09-350-641C-1420
; Sequence 1420, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-641C-1420

Query Match      23.5%; Score 23; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      12 NGTF 15

RESULT 27
US-09-350-841A-1420
; Sequence 1420, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-841A-1420

Query Match      23.5%; Score 23; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      33 NGTF 36
Db      12 NGTF 15

RESULT 28
US-09-082-279B-1419
; Sequence 1419, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT FILING DATE: 2001-04-13
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1419
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1419

Query Match      23.5%; Score 23; DB 3; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      14 NGTF 17

RESULT 29
US-09-315-304B-1419
; Sequence 1419, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1419
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1419

Query Match      23.5%; Score 23; DB 3; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      14 NGTF 17

RESULT 30
US-09-834-784-1419
; Sequence 1419, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; CURRENT FILING DATE: 2001-04-13
; PRIOR FILING DATE: 09/082,279
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1419
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-834-784-1419

Query Match      23.5%; Score 23; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      14 NGTF 17

RESULT 31
US-09-515-965A-1419
; Sequence 1419, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515,965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR FILING DATE: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR FILING DATE: 09/082,279
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1419
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-515-965A-1419

Query Match      23.5%; Score 23; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
```


Db 14 NGTF 17

RESULT 40
US-09-350-641C-1242
; Sequence 1242, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; TITLE OF INVENTION: PROPERTIES
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1242
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-350-641C-1242

Query Match 23.5%; Score 23; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||

Db 14 NGTF 17

Search completed: February 24, 2005, 15:25:30
Job time : 30.2276 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:23:51 ; Search time 81.0406 Seconds
(without alignments)
226.127 Million cell updates/sec

Title: US-09-936-956-1

Perfect score: 98

Sequence: 1 XGXXXXXXXXXXXXXXXXXAX.....XXXXXXXXXXXXXXXXXXXXXG 56

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	25.5	260	16	US-10-652-870-248
2	25	25.5	260	16	US-10-652-870-252
3	25	25.5	260	16	US-10-652-870-250
4	25	25.5	1026	9	US-09-379-931-7
5	25	25.5	1026	14	US-10-233-597-7
6	25	25.5	1026	17	US-10-478-676-1
7	23	23.5	9	16	US-10-657-022-194
8	23	23.5	10	15	US-10-245-871-305
9	23	23.5	10	15	US-10-296-317-48
10	23	23.5	10	15	US-10-253-286-305
11	23	23.5	12	10	US-09-991-225-23
12	23	23.5	12	15	US-10-369-405-23
13	23	23.5	14	10	US-09-991-225-42

14	23	23.5	14	10	US-09-991-225-69	Sequence 69, Appl
15	23	23.5	14	15	US-10-369-405-42	Sequence 42, Appl
16	23	23.5	14	15	US-10-369-405-69	Sequence 69, Appl
17	23	23.5	15	10	US-09-991-225-60	Sequence 60, Appl
18	23	23.5	15	10	US-09-991-225-72	Sequence 72, Appl
19	23	23.5	15	14	US-10-225-567A-2046	Sequence 2046, Ap
20	23	23.5	15	15	US-10-369-405-60	Sequence 60, Appl
21	23	23.5	15	15	US-10-369-405-72	Sequence 72, Appl
22	23	23.5	18	10	US-09-962-756-606	Sequence 606, App
23	23	23.5	18	15	US-10-253-471-606	Sequence 606, App
24	23	23.5	18	15	US-10-253-493-606	Sequence 606, App
25	23	23.5	24	15	US-10-245-871-312	Sequence 312, App
26	23	23.5	24	15	US-10-253-286-312	Sequence 312, App
27	23	23.5	26	15	US-10-296-317-67	Sequence 67, Appl
28	23	23.5	28	15	US-10-296-317-59	Sequence 59, Appl
29	23	23.5	30	15	US-10-424-599-152465	Sequence 152465,
30	23	23.5	30	15	US-10-424-599-168021	Sequence 168021,
31	23	23.5	31	9	US-09-209-799D-22	Sequence 22, Appl
32	23	23.5	31	9	US-09-736-960-5	Sequence 5, Appl
33	23	23.5	31	10	US-09-997-792-22	Sequence 22, Appl
34	23	23.5	37	9	US-09-995-494-113	Sequence 113, App
35	23	23.5	38	14	US-10-351-641-1420	Sequence 1420, Ap
36	23	23.5	39	14	US-10-351-641-1419	Sequence 1419, Ap
37	23	23.5	40	14	US-10-351-641-1242	Sequence 1242, Ap
38	23	23.5	40	14	US-10-351-641-1747	Sequence 1747, Ap
39	23	23.5	40	14	US-10-351-641-1748	Sequence 1748, Ap
40	23	23.5	41	9	US-09-912-628-9	Sequence 9, Appli
41	23	23.5	41	14	US-10-116-166-7	Sequence 7, Appli
42	23	23.5	41	15	US-10-425-114-44378	Sequence 44378, A
43	23	23.5	42	14	US-10-351-641-1243	Sequence 1243, Ap
44	23	23.5	42	16	US-10-437-963-180144	Sequence 180144,
45	23	23.5	43	15	US-10-424-599-238104	Sequence 238104,

ALIGNMENTS

RESULT 1

US-10-652-870-248
; Sequence 248, Application US/10652870
; Publication No. US20040167068A1
; GENERAL INFORMATION:
; APPLICANT: Zlotnick, Gary
; APPLICANT: Fletcher, Leah
; APPLICANT: John, Farley
; APPLICANT: Bernfield, Liesel
; APPLICANT: Zagursky, Robert
; APPLICANT: Metcalf, Benjamin
; TITLE OF INVENTION: Novel Immunogenic Compositions for the Prevention and Treatment o
; FILE REFERENCE: 38523.000026
; CURRENT APPLICATION NUMBER: US/10/652,870
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/652,870
; PRIOR FILING DATE: 2003-09-02
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 248
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-652-870-248

Query Match 25.5%; Score 25; DB 16; Length 260;
Best Local Similarity 25.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 33 NGTFFXXXXXXXXXXXXGDXXXG 56

Db 48 NGTTLTSAQGAETKYGNGDSLNTG 71

RESULT 2

```
US-10-652-870-252
; Sequence 252, Application US/10652870
; Publication No. US20040167068A1
; GENERAL INFORMATION:
; APPLICANT: Zlotnick, Gary
; APPLICANT: Fletcher, Leah
; APPLICANT: John, Farley
; APPLICANT: Bernfield, Liesel
; APPLICANT: Zagursky, Robert
; APPLICANT: Metcalf, Benjamin
; TITLE OF INVENTION: Novel Immunogenic Compositions for the Prevention and Treatment of Meningococcal Disease
; FILE REFERENCE: 38523.000026
; CURRENT APPLICATION NUMBER: US/10/652,870
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/652,870
; PRIOR FILING DATE: 2003-09-02
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 252
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-652-870-252

Query Match      25.5%; Score 25; DB 16; Length 260;
Best Local Similarity 25.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXXXXXXGDXXXG 56
DB 48 NGTLTSAQGAERTYGNGLSINTG 71

RESULT 3
US-10-652-870-250
; Sequence 250, Application US/10652870
; Publication No. US20040167068A1
; GENERAL INFORMATION:
; APPLICANT: Zlotnick, Gary
; APPLICANT: Fletcher, Leah
; APPLICANT: John, Farley
; APPLICANT: Bernfield, Liesel
; APPLICANT: Zagursky, Robert
; APPLICANT: Metcalf, Benjamin
; TITLE OF INVENTION: Novel Immunogenic Compositions for the Prevention and Treatment of Meningococcal Disease
; FILE REFERENCE: 38523.000026
; CURRENT APPLICATION NUMBER: US/10/652,870
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/652,870
; PRIOR FILING DATE: 2003-09-02
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 250
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-652-870-250

Query Match      25.5%; Score 25; DB 16; Length 261;
Best Local Similarity 25.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXXXXXXGDXXXG 56
DB 49 NGTLTSAQGAERTYGNGLSINTG 72

RESULT 4
US-09-379-931-7
; Sequence 7, Application US/09379931
; Patent No. US20020009792A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H.
; APPLICANT: No. US20020009792A1e1lini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CAULOBACTER CRESCENTUS
; FILE REFERENCE: 08106/002003
; CURRENT APPLICATION NUMBER: US/09/379,931
; CURRENT FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: US 08/614,377
; PRIOR FILING DATE: 1996-03-12
; PRIOR APPLICATION NUMBER: US 08/194,290
; PRIOR FILING DATE: 1994-02-09
; PRIOR APPLICATION NUMBER: US 07/895,367
; PRIOR FILING DATE: 1992-06-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-09-379-931-7

Query Match      25.5%; Score 25; DB 9; Length 1026;
Best Local Similarity 25.0%; Pred. No. 4.5e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXXXXXXGDXXXG 56
DB 267 NDTFVAGEVAGAAATLTVGDTLSGG 290

RESULT 5
US-10-223-597-7
; Sequence 7, Application US/10223597
; Publication No. US20030135037A1
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H.
; APPLICANT: No. US20030135037A1e1lini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CAULOBACTER CRESCENTUS
; FILE REFERENCE: 08106/002003
; CURRENT APPLICATION NUMBER: US/10/223,597
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US/09/379,931
; PRIOR FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: US 08/614,377
; PRIOR FILING DATE: 1996-03-12
; PRIOR APPLICATION NUMBER: US 08/194,290
; PRIOR FILING DATE: 1994-02-09
; PRIOR APPLICATION NUMBER: US 07/895,367
; PRIOR FILING DATE: 1992-06-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-223-597-7

Query Match      25.5%; Score 25; DB 14; Length 1026;
Best Local Similarity 25.0%; Pred. No. 4.5e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXXXXXXGDXXXG 56
DB 267 NDTFVAGEVAGAAATLTVGDTLSGG 290

RESULT 6
US-10-478-676-1
; Sequence 1, Application US/10478676
; Publication No. US20050032194A1
; GENERAL INFORMATION:
```



```
; APPLICANT: SMIT, John
; APPLICANT: NOMELELLI, John, F.
; APPLICANT: BINGLES, Wade, H.
; TITLE OF INVENTION: PROTEASE DEFICIENT CAULOBACTER HOST CELLS
; FILE REFERENCE: 32235-199048
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/10/478,676
; PRIOR FILING DATE: 2002-05-22 WO PCT/CA02/00722
; PRIOR APPLICATION NUMBER: CA 2,347,657
; PRIOR FILING DATE: 2001-05-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-478-676-1

Query Match      25.5%; Score 25; DB 17; Length 1026;
Best Local Similarity 25.0%; Pred. No. 4.5e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXXGDXG 56
Db 267 NDTFVAGEVAGATLTVDLTSGG 290

RESULT 7
US-10-657-022-194
; Sequence 194, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPTOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homosapiens
US-10-657-022-194

Query Match      23.5%; Score 23; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 3 NGTF 6

RESULT 8
US-10-245-871-305
; Sequence 305, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17

Qy 33 NGTF 36
Db 3 NGTF 6

RESULT 9
US-10-296-317-48
; Sequence 48, Application US/10296317
; Publication No. US20040057968A1
; GENERAL INFORMATION:
; APPLICANT: CEL-Sci Corp
; APPLICANT: Zimmerman, Daniel S
; APPLICANT: Sarin, Prem S
; TITLE OF INVENTION: T CELL BINDING LIGAND PEPTIDES, PEPTIDE
; FILE REFERENCE: CS-112
; CURRENT APPLICATION NUMBER: US/10/296,317
; CURRENT FILING DATE: 2002-11-23
; PRIOR APPLICATION NUMBER: US 60/206548
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: PCT/US07/16793
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CEA Peptide C3
US-10-296-317-48

Query Match      23.5%; Score 23; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 7 NGTF 10

RESULT 10
US-10-253-286-305
; Sequence 305, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 305
; LENGTH: 10
```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-253-286-305

Query Match 23.5%; Score 23; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|
|
|
|
Db 7 NGTF 10

RESULT 11

US-09-991-225-23
; Sequence 23, Application US/09991225
; Publication No. US20030153063A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, EXPRESSED HI
; FILE REFERENCE: D0075.NP
; CURRENT APPLICATION NUMBER: US/09/991,225
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; SOFTWARE: PatentIn version 3.0
; NUMBER OF SEQ ID NOS: 81
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-991-225-23

Query Match 23.5%; Score 23; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|
|
|
|
Db 4 NGTF 7

RESULT 12

US-10-369-405-23
; Sequence 23, Application US/10369405
; Publication No. US20030224400A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, AND VARIANTS
; FILE REFERENCE: D0075A CIP
; CURRENT APPLICATION NUMBER: US/10/369,405
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: U.S. 09/991,225
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: U.S. 60/305,818
; PRIOR FILING DATE: 2001-07-16
; SOFTWARE: PatentIn version 3.2
; NUMBER OF SEQ ID NOS: 94
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-369-405-23

Query Match 23.5%; Score 23; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|
|
|
|
Db 4 NGTF 7

RESULT 13

US-09-991-225-42
; Sequence 42, Application US/09991225
; Publication No. US20030153063A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, EXPRESSED HI
; FILE REFERENCE: D0075.NP
; CURRENT APPLICATION NUMBER: US/09/991,225
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-991-225-42

Query Match 23.5%; Score 23; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|
|
|
|
Db 6 NGTF 9

RESULT 14

US-09-991-225-69
; Sequence 69, Application US/09991225
; Publication No. US20030153063A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, EXPRESSED HI
; FILE REFERENCE: D0075.NP
; CURRENT APPLICATION NUMBER: US/09/991,225
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-991-225-69

Query Match 23.5%; Score 23; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 11 NGTF 14

RESULT 15
US-10-369-405-42
; Sequence 42, Application US/10369405
; Publication No. US20030224400A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, AND VARIANTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: D0075A CIP
; CURRENT APPLICATION NUMBER: US/10/369,405
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: U.S. 09/991,225
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: U.S. 60/305,818
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-405-42

Query Match 23.5%; Score 23; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 6 NGTF 9

RESULT 16
US-10-369-405-69
; Sequence 69, Application US/10369405
; Publication No. US20030224400A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, AND VARIANTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: D0075A CIP
; CURRENT APPLICATION NUMBER: US/10/369,405
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: U.S. 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: U.S. 09/991,225
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: U.S. 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: U.S. 60/305,818
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-405-69

Query Match 23.5%; Score 23; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36

Db 11 NGTF 14
|||||

RESULT 17
US-09-991-225-60
; Sequence 60, Application US/09991225
; Publication No. US20030153063A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, EXPRESSED HI
; TITLE OF INVENTION: HEART AND VARIANTS THEREOF
; FILE REFERENCE: D0075.NP
; CURRENT APPLICATION NUMBER: US/09/991,225
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 60
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-991-225-60

Query Match 23.5%; Score 23; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 4 NGTF 7

RESULT 18
US-09-991-225-72
; Sequence 72, Application US/09991225
; Publication No. US20030153063A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV11, EXPRESSED HI
; TITLE OF INVENTION: HEART AND VARIANTS THEREOF
; FILE REFERENCE: D0075.NP
; CURRENT APPLICATION NUMBER: US/09/991,225
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/249,613
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/257,611
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/305,818
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-991-225-72

Query Match 23.5%; Score 23; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||||
Db 4 NGTF 7

RESULT 19

```
US-10-225-567A-2046
; Sequence 2046, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2046
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-2046

Query Match      23.5%; Score 23; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      |||||
        4 NGTF 7

RESULT 20
US-10-369-405-60
; Sequence 60, Application US/10369405
; Publication No. US20030224400A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPREMY11, AND VARIANTS
; FILE REFERENCE: D0075A CIP
; CURRENT APPLICATION NUMBER: US/10/369,405
; CURRENT FILING DATE: 2003-02-14
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2001-11-16
; PRIOR FILING DATE: 2001-11-16
; PRIOR FILING DATE: 2001-11-16
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-405-60

Query Match      23.5%; Score 23; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      |||||
        4 NGTF 7

RESULT 21
US-10-369-405-72
; Sequence 72, Application US/10369405
; Publication No. US20030224400A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
```

```
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPREMY11, AND VARIANTS
; FILE REFERENCE: D0075A CIP
; CURRENT APPLICATION NUMBER: US/10/369,405
; CURRENT FILING DATE: 2003-02-14
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2001-11-16
; PRIOR FILING DATE: 2001-11-16
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-405-72

Query Match      23.5%; Score 23; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      |||||
        4 NGTF 7

RESULT 22
US-09-962-756-606
; Sequence 606, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAEFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 606
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-606

Query Match      23.5%; Score 23; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
Db      |||||
        5 NGTF 8

RESULT 23
```

```
US-10-253-471-606
; Sequence 606, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTIA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 606
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-606

Query Match      23.5%; Score 23; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
DB      5 NGTF 8

RESULT 24
US-10-253-493-606
; Sequence 606, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTIA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 606
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-493-606

Query Match      23.5%; Score 23; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
DB      5 NGTF 8

US-10-253-471-606
; Sequence 312, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 312
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Ii-key/MHC Class II/MHC Class I CEA hybrid
; OTHER INFORMATION: peptide
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: a-aminovaleric acid
; FEATURE:
; OTHER INFORMATION: C-term amidated
US-10-245-871-312

Query Match      23.5%; Score 23; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      33 NGTF 36
DB      21 NGTF 24

RESULT 26
US-10-253-286-312
; Sequence 312, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 312
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Ii-key/MHC Class II/MHC Class I CEA hybrid
; OTHER INFORMATION: peptide
; NAME/KEY: MOD_RES
; LOCATION: (5)
; OTHER INFORMATION: a-aminovaleric acid
; FEATURE:
; OTHER INFORMATION: C-term amidated
US-10-253-286-312
```

Query Match 23.5%; Score 23; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 21 NGTF 24

RESULT 27
US-10-296-317-67
; Sequence 67, Application US/10296317
; Publication No. US20040057968A1
; GENERAL INFORMATION:
; APPLICANT: CEI-Sci Corp
; APPLICANT: Sarin, Prem S
; APPLICANT: Zimmerman, Daniel S
; TITLE OF INVENTION: T CELL BINDING LIGAND PEPTIDES, PEPTIDE
; FILE REFERENCE: CS-112
; CURRENT APPLICATION NUMBER: US/10/296,317
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/206548
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: PCT/US07/16793
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 67
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide construct
US-10-296-317-67

Query Match 23.5%; Score 23; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 23 NGTF 26

RESULT 28
US-10-296-317-59
; Sequence 59, Application US/10296317
; Publication No. US20040057968A1
; GENERAL INFORMATION:
; APPLICANT: CEI-Sci Corp
; APPLICANT: Zimmerman, Daniel S
; APPLICANT: Sarin, Prem S
; TITLE OF INVENTION: T CELL BINDING LIGAND PEPTIDES, PEPTIDE
; FILE REFERENCE: CS-112
; CURRENT APPLICATION NUMBER: US/10/296,317
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/206548
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: PCT/US07/16793
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 59
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide construct
US-10-296-317-59

Query Match 23.5%; Score 23; DB 15; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 25 NGTF 28

RESULT 29
US-10-424-599-152465
; Sequence 152465, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152465
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108700C.1.pap
US-10-424-599-152465

Query Match 23.5%; Score 23; DB 15; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 17 NGTF 20

RESULT 30
US-10-424-599-168021
; Sequence 168021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 168021
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_122739C.1.pap
US-10-424-599-168021

Query Match 23.5%; Score 23; DB 15; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
Db 3 NGTF 6

RESULT 31
US-09-209-799D-22
; Sequence 22, Application US/09209799D

```
; Publication No. US20010014666A1
; GENERAL INFORMATION:
; APPLICANT: Hermeling, Ronald
; APPLICANT: Hoffmann, James
; APPLICANT: Narasimhan, Chakravarthy
; TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
; FILE REFERENCE: X-10242
; CURRENT APPLICATION NUMBER: US/09/209,799D
; CURRENT FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-209-799D-22

Query Match      23.5%; Score 23; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
      ||||
Db      3 NGTF 6

RESULT 32
US-09-736-960-5
; Sequence 5, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn ver. 2.1
```

```
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: preliminary human CLASP-5
US-09-736-960-5

Query Match      23.5%; Score 23; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
      ||||
Db      24 NGTF 27

RESULT 33
US-09-997-792-22
; Sequence 22, Application US/09997792
; Publication No. US20030045464A1
; GENERAL INFORMATION:
; APPLICANT: Hermeling, Ronald
; APPLICANT: Hoffmann, James
; APPLICANT: Narasimhan, Chakravarthy
; TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
; FILE REFERENCE: X-10242
; CURRENT APPLICATION NUMBER: US/09/997,792
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-997-792-22

Query Match      23.5%; Score 23; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
      ||||
Db      3 NGTF 6

RESULT 34
US-09-995-494-113
; Sequence 113, Application US/09995494
; Patent No. US20020127578A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0293
; CURRENT APPLICATION NUMBER: US/09/995,494
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,176
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-995-494-113
```

```
Query Match      23.5%; Score 23; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      17 NGTF 20

RESULT 35
US-10-351-641-1420
; Sequence 1420, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1420
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1420

Query Match      23.5%; Score 23; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      12 NGTF 15

RESULT 36
US-10-351-641-1419
; Sequence 1419, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1419
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1419

Query Match      23.5%; Score 23; DB 14; Length 39;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      14 NGTF 17

RESULT 37
US-10-351-641-1242
; Sequence 1242, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1242
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1242

Query Match      23.5%; Score 23; DB 14; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      14 NGTF 17

RESULT 38
US-10-351-641-1747
; Sequence 1747, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1747
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1747
```



```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1747
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1747

Query Match      23.5%; Score 23; DB 14; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      14 NGTF 17

RESULT 39
US-10-351-641-1748
; Sequence 1748, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1748
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 4, 8
; OTHER INFORMATION: Xaa = Abu (aminobutyric acid)
US-10-351-641-1748

Query Match      23.5%; Score 23; DB 14; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      14 NGTF 17

RESULT 40
US-09-912-628-9
; Sequence 9, Application US/09912628
; Patent No. US20020160491A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Human Serpin Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT001P2
; CURRENT APPLICATION NUMBER: US/09/912,628
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: PCT/US01/02484
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,769
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: PCT/US00/05082
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-912-628-9

Query Match      23.5%; Score 23; DB 9; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      8 NGTF 11

Search completed: February 24, 2005, 15:42:01
Job time : 83.0406 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:12:40 ; Search time 21.8537 Seconds
(without alignments)
246.555 Million cell updates/sec

Title: US-09-936-956-1
Perfect score: 98
Sequence: 1 XGXXXXXXXXXXXXXAX.....XXXXXXXXXXXXGXXXXXG 56

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	28.6	865	1 D70986	probable ABC trans
2	25	25.5	220	2 T34986	probable secreted
3	25	25.5	312	2 A31846	130K paracrystalli
4	25	25.5	1026	2 A48995	paracrystalline su
5	25	25.5	1073	2 C87374	S-layer protein Rg
6	24	24.5	435	2 A61028	prepilin imported
7	24	24.5	698	2 D81832	transferrin-bindin
8	23	23.5	42	2 E81911	hypothetical prote
9	23	23.5	49	2 E85731	Rhs element associ
10	23	23.5	57	2 D96953	hypothetical prote
11	23	23.5	83	2 T40973	cytochrome c oxida
12	23	23.5	84	2 A96525	protein rin15.13 [
13	23	23.5	102	2 D26300	alpha-1-acid glyco
14	23	23.5	102	2 A82829	colicin V precursor
15	23	23.5	102	2 B82829	colicin V precursor
16	23	23.5	110	2 S43147	Ig upailon chain -
17	23	23.5	111	2 E89223	hypothetical prote
18	23	23.5	112	2 A03596	hypothetical cytos
19	23	23.5	116	2 S44709	opacity protein op
20	23	23.5	119	2 AD0935	probable membrane
21	23	23.5	119	2 D86088	hypothetical prote
22	23	23.5	119	2 G31240	hypothetical prote
23	23	23.5	119	2 G65203	Y1Jp protein - Esc
24	23	23.5	120	2 S21565	hypothetical prote
25	23	23.5	121	2 E97103	hypothetical prote
26	23	23.5	123	2 AD1123	B. subtilis YhdG p
27	23	23.5	132	2 AF1483	B. subtilis YhdG p
28	23	23.5	134	2 T36750	probable gntR-fam
29	23	23.5	135	2 D84339	hypothetical prote

30	23	23.5	138	2 A99362	hypothetical prote
31	23	23.5	138	2 T24916	hypothetical prote
32	23	23.5	141	2 S67149	hypothetical prote
33	23	23.5	143	2 T49689	hypothetical prote
34	23	23.5	143	2 G91103	hypothetical prote
35	23	23.5	143	2 E71041	hypothetical prote
36	23	23.5	143	2 B85949	hypothetical prote
37	23	23.5	144	2 T33494	hypothetical prote
38	23	23.5	145	2 S72492	probable proteinase
39	23	23.5	147	2 A69883	conserved hypothet
40	23	23.5	151	2 F69744	hypothetical prote
41	23	23.5	152	2 T21761	hypothetical prote
42	23	23.5	156	2 T02166	cysteine proteinase
43	23	23.5	156	2 A32795	T-cell translocati
44	23	23.5	157	2 H97061	hypothetical prote
45	23	23.5	158	2 T07597	proteinase inhibit

ALIGNMENTS

RESULT 1

D70986
probable ABC transporter Rv1747 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: D70986
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70986
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-865 <COL>
A:Cross-references: UNIPROT:O65934; GB:Z95890; GB:AL123456; NID:g3242245; PIDN:CAB09333
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1747
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter Rv1747; ATP-binding
C:Keywords: ATP
F:335-528/Domain: ATP-binding cassette homology <ABC>

Query Match 28.6%; Score 28; DB 1; Length 865;
Best Local Similarity 29.2%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXGDXGXXXXXG 56
|||
Db 270 NGTFVNGARVDAALLHGDGVVTIG 293

RESULT 2

T34986
probable secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34986
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, August 1998
A:Reference number: Z21550
A:Accession: T34986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-220 <OLI>
A:Cross-references: UNIPROT:O86667; EMBL:AL031182; PIDN:CAA20163.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOE:DB:SC4A2.11c

Query Match 25.5%; Score 25; DB 2; Length 220;
Best Local Similarity 26.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXG 51
Db 158 HGTFTTLEGGSTVTTSGSGD 176

RESULT 3
A31846
130K paracrystalline protein - Caulobacter crescentus (fragment)
C:Species: Caulobacter crescentus
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Mar-1998
C:Accession: A31846
R:Fisher, J.A.; Smit, J.; Agabian, N.
J. Bacteriol. 170, 4706-4713, 1988
A:Title: Transcriptional analysis of the major surface array gene of Caulobacter crescentus
A:Reference number: A31846; MUID:89008089; PMID:3049545
A:Accession: A31846
A:Molecule type: DNA
A:Residues: 1-312 <FIS>
A:Cross-references: GB:M22663; NID:g556033; PID:g556034

Query Match 25.5%; Score 25; DB 2; Length 312;
Best Local Similarity 25.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXG 56
Db 265 NDTFVAGEVAGAAATLTVGDTLSGG 288

RESULT 4
A48995
Paracrystalline surface layer protein RsaA - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A48995
R:Gilchrist, A.; Fisher, J.A.; Smit, J.
Can. J. Microbiol. 38, 193-202, 1992
A:Title: Nucleotide sequence analysis of the gene encoding the Caulobacter crescentus p
A:Reference number: A48995; MUID:93007489; PMID:1393820
A:Accession: A48995
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-1026 <GIL>
A:Cross-references: UNIPROT:P35828; GB:AF062345; GB:M22663; GB:M84760; NID:g6064104; PID
A:Experimental source: CB15A, ATCC 19089
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:116173, NCBIP:116174)

Query Match 25.5%; Score 25; DB 2; Length 1026;
Best Local Similarity 25.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXG 56
Db 267 NDTFVAGEVAGAAATLTVGDTLSGG 290

RESULT 5
C87374
S-layer protein RsaA [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87374
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87374
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1073 <STO>
A:Cross-references: GB:AE005673; NID:g13422297; PIDN:AAK22991.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1007

Query Match 25.5%; Score 25; DB 2; Length 1073;
Best Local Similarity 25.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXG 56
Db 314 NDTFVAGEVAGAAATLTVGDTLSGG 337

RESULT 6
AG1028
prepilin [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG1028
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG1028
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-435 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD09328.1; PID:g16505328; GSPDB:GN00176
C:Genetics:
A:Gene: pilV
C:Superfamily: shufflon

Query Match 24.5%; Score 24; DB 2; Length 435;
Best Local Similarity 25.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXG 56
Db 238 NGNFGVSLVSNGPITAGDIRSTG 261

RESULT 7
D81832
transferrin-binding protein B NMA2025 [imported] - Neisseria meningitidis (strain Z2491)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: D81832
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81832
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-698 <PAR>
A:Cross-references: UNIPROT:O68937; GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85244
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ttpB; NMA2025
C:Superfamily: bacterial pathogen transferrin-binding protein; tonB-dependent receptor an

Query Match 24.5%; Score 24; DB 2; Length 698;
Best Local Similarity 25.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

A;Gene: TIM15.13
A;Map position: 1

Query Match 23.5%; Score 23; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 19 NGTF 22

RESULT 13

B26300
alpha-1-acid glycoprotein (clone pMAGP3) - mouse (fragment)
N;Alternate names: orosomucoid
C;Species: Mus musculus (house mouse)
C;Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 09-Jul-2004
C;Accession: B26300
R;Cooper, R.; Papaconstantinou, J.
J. Biol. Chem. 261, 1849-1853, 1986
A;Title: Evidence for the existence of multiple alpha-1-acid glycoprotein genes in the m
A;Reference number: A26300; MUID:86111861; PMID:3003086
A;Accession: B26300
A;Molecule type: mRNA
A;Residues: 1-102 <COO>
A;Cross-references: UNIPROT:Q91X23
A;Experimental source: strain SWR/J
C;Comment: At least 2 genes coding for different forms of alpha-1-AGP appear to be pres
C;Comment: Alpha-1-AGP is synthesized in the liver, the synthesis being controlled by gl
C;Genetics: AGP
A;Gene: AGP
C;Superfamily: lipocalin; lipocalin homology
C;Keywords: acute phase; glycoprotein; liver; plasma
F;1-101/Domain: lipocalin homology (fragment) <LIP>
F;8-101/Disulfide bonds: #status predicted
F;11,21/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 23.5%; Score 23; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 21 NGTF 24

RESULT 14

A82829
colicin V precursor XF0262 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82829
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-102 <SIM>
A;Cross-references: UNIPROT:Q9PGN6; GB:AE003879; GB:AE003849; NID:g9105067; PIDN:AAF8307
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; de Oliveira, R.C.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0262

Query Match 23.5%; Score 23; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 97 NGTF 100

RESULT 15

B82829
colicin V precursor XF0263 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: B82829
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-102 <SIM>
A;Cross-references: UNIPROT:Q9PGN5; GB:AE003879; GB:AE003849; NID:g9105067; PIDN:AAF8307
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

Query Match 23.5%; Score 23; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 97 NGTF 100

RESULT 16

S43147
IG upsilon chain - duck (fragment)
C;Species: Anas platyrhynchos (domestic duck)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 11-Jan-2000
C;Accession: S43147
R;Magor, K.E.; Higgins, D.A.; Middleton, D.I.; Warr, G.W.
submitted to the EMBL Data Library, March 1994
A;Description: Evidence from duck immunoglobulin genes that IgY is the common ancestor of
A;Reference number: S43145
A;Accession: S43147
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-110 <MAG>
 A;Cross-references: EMBL:X78355; NID:g468612; PID:g468613
 C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 23.5%; Score 23; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 67 NGTF 70

RESULT 17

E69223
 hypothetical protein MTH922 - Methanobacterium thermoautotrophicum (strain Delta H)
 C;Species: Methanobacterium thermoautotrophicum
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C;Accession: E69223
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcu
 A;Reference number: A69000; MUID:98037514; PMID:9371463
 A;Accession: E69223
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-111 <MTH>
 A;Cross-references: UNIPROT:O27005; GB:AE000867; GB:AE000666; NID:g2622009; PIDN:AAB8542
 A;Experimental source: strain Delta H
 C;Genetics:
 A;Gene: MTH922
 A;Start codon: TTG
 C;Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH922

Query Match 23.5%; Score 23; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 72 NGTF 75

RESULT 18

AD3596
 hypothetical cytosolic protein BMEII0693 [imported] - Brucella melitensis (strain 16M)
 C;Species: Brucella melitensis
 C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C;Accession: AD3596
 R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
 proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
 A;Reference number: AD3252; PMID:1175668

A;Accession: AD3596
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-112 <KUR>
 A;Cross-references: UNIPROT:Q8VC39; GB:AE008918; PIDN:AAL53935.1; PID:g17984880; GSPDB:G
 A;Experimental source: strain 16M
 C;Genetics:
 A;Gene: BMEII0693
 A;Map position: II

Query Match 23.5%; Score 23; DB 2; Length 112;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 101 NGTF 104

RESULT 19

S44709
 opacity protein opaB - Neisseria meningitidis (isolate Z1213) (fragment)
 C;Species: Neisseria meningitidis
 A;Variety: isolate Z1213
 C;Date: 13-Jan-1995 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
 C;Accession: S44709
 R;Hobbs, M.M.; Seiler, A.; Achtman, M.; Cannon, J.G.
 Mol. Microbiol. 12, 171-180, 1994

A;Title: Microevolution within a clonal population of pathogenic bacteria: recombination
 A;Reference number: S44705; MUID:94335637; PMID:7520117
 A;Accession: S44709

A;Molecule type: DNA

A;Residues: 1-116 <HOB>
 A;Cross-references: UNIPROT:Q51089; EMBL:U03409; NID:g432174; PIDN:AAA61550.1; PID:g4321
 A;Experimental source: isolate Z1213; strain serogroup A; clone pFLOB2510
 A;Note: expression of opacity proteins is regulated by the number of translated repeat e
 of repeats place the start codon in frame with the rest of the protein
 C;Genetics:
 A;Gene: opaB

C;Superfamily: opacity protein
 C;Keywords: cell surface component; transmembrane protein
 F;1-37/Domain: extracellular #status predicted <EXT2>
 F;7-32/Region: hypervariable region HV1
 F;38-52/Domain: transmembrane #status predicted <TM4>
 F;61-71/Domain: transmembrane #status predicted <TM5>
 F;77-116/Region: hypervariable region HV2

Query Match 23.5%; Score 23; DB 2; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 35 NGTF 38

RESULT 20

AD0935
 probable membrane protein STY3746 [imported] - Salmonella enterica subsp. enterica serov
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AD0935

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AD0935

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-119 <PAR>

A;Cross-references: PIDN:CAD09502.1; PID:g16504619; GSPDB:GN00176

C;Genetics:

A;Gene: STV3746

Query Match 23.5%; Score 23; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 22 NGTF 25

RESULT 21

D86088
 hypothetical protein yijD [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D86088
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D86088
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-119 <STO>
A;Cross-references: UNIPROT:P27308; GB:AE005174; NID:G12518891; PIDN:AAG59168.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yjJd

Query Match 23.5%; Score 23; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 22 NGTF 25

RESULT 22
G91240
hypothetical protein EC4895 [imported] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: G91240
R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G91240
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-119 <HAY>
A;Cross-references: UNIPROT:P27308; GB:BA000007; PIDN:BA838318.1; PID:G13364371; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC4895

Query Match 23.5%; Score 23; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 22 NGTF 25

RESULT 23
G65203
yJd protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: G65203
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-119 <BLAT>
A;Cross-references: UNIPROT:P27308; GB:AE000470; GB:U000096; NID:G2367332; PIDN:AAC76946.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:

A;Gene: yjJd

Query Match 23.5%; Score 23; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 22 NGTF 25

RESULT 24
S21565
hypothetical protein B - Escherichia coli
C;Species: Escherichia coli
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S21565
R;Gustafsson, C.
submitted to the EMBL Data Library, May 1992
A;Reference number: S21563
A;Accession: S21565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-120 <GUS>
A;Cross-references: UNIPROT:P27308; EMBL:X66026; NID:G43236; PIDN:CAA46824.1; PID:G43239

Query Match 23.5%; Score 23; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 23 NGTF 26

RESULT 25
E97103
hypothetical protein CAC1650 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97103
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-121 <KUR>
A;Cross-references: UNIPROT:Q971J1; GB:AE001437; PIDN:AAK79616.1; PID:G15024609; GSPDB:GN
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1650

Query Match 23.5%; Score 23; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
|||
Db 36 NGTF 39

RESULT 26
AD1123
B. subtilis yhdG protein homolog lmo0387 [imported] - Listeria monocytogenes (strain EGO
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD1123
R;Glasier, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Feihl, H.,
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1123
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <GLA>
A:Cross-references: UNIPROT:Q8Y9Y0; GB:NC_003210; PIDN:CAC98466.1; PID:gl6409765; GSPDB:
C:Genetics:
A:Gene: lmo0387
C:Superfamily: Bacillus subtilis hypothetical protein ydhg

Query Match 23.5%; Score 23; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||
DB 48 NGTF 51

RESULT 27
AF1483
B. subtilis YhdG protein homolog lin0405 [imported] - *Listeria innocua* (strain Clip11262
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF1483
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Feihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1483
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-132 <GLA>
A:Cross-references: UNIPROT:Q92EQ3; GB:AL592022; PIDN:CAC95638.1; PID:gl6412834; GSPDB:G
A:Experimental source: strain clip11262
C:Genetics:
A:Gene: lin0405
C:Superfamily: Bacillus subtilis hypothetical protein ydhg

Query Match 23.5%; Score 23; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||
DB 48 NGTF 51

RESULT 28
T36750
probable gntR-family transcription regulator - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36750
R:Saunders, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: 221613
A:Accession: T36750
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-134 <SAU>
A:Cross-references: UNIPROT:Q92A2; EMBL:AL096849; PIDN:CAB50943.1; GSPDB:GN00070; SCOE
A:Experimental source: strain A3(2)
C:Genetics:

A:Gene: SCOEDB:SC111.17C

Query Match 23.5%; Score 23; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||
DB 85 NGTF 88

RESULT 29
D64339
hypothetical protein MJ0315 - *Methanococcus jannaschii*
C:Species: *Methanococcus jannaschii*
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: D64339
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: D64339
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-135 <BLU>
A:Cross-references: GB:U67486; GB:L77117; NID:gi591031; PIDN:AAB98301.1; PID:gi591035; T
C:Genetics:
A:Map position: REV294859-294452
A:Start codon: TTG
C:Superfamily: conserved hypothetical protein MJ0315

Query Match 23.5%; Score 23; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||
DB 129 NGTF 132

RESULT 30
A99362
hypothetical protein SSO1961 [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: A99362
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: A99362
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-138 <KUR>
A:Cross-references: UNIPROT:Q97WZ9; GB:AE006641; NID:gi13815234; PIDN:AAK42152.1; GSPDB:G
C:Genetics:
A:Gene: SSO1961

Query Match 23.5%; Score 23; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NGTF 36
|||
DB 127 NGTF 130

RESULT 31

T24916
 hypothetical protein T14G10.4 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T24916
 R;Wild, A.
 submitted to the EMBL Data Library, January 1996
 A;Reference number: Z19954
 A;Accession: T24916
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-138 <WIL>
 A;Cross-references: UNIPROT:Q22499; EMBL:Z68880; PIDN:CAA93096.1; GSPDB:GN00022; CESP:TL
 A;Experimental source: clone T14G10
 C;Genetics:
 A;Gene: CESP:T14G10.4
 A;Map position: 4
 A;Introns: 52/3; 108/2

Query Match 23.5%; Score 23; DB 2; Length 138;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 74 NGTF 77

RESULT 32
 S67149
 hypothetical protein YOR252w - yeast (*Saccharomyces cerevisiae*)
 N;Alternate names: hypothetical protein O5310
 C;Species: *Saccharomyces cerevisiae*
 C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
 C;Accession: S67149
 R;Jauniaux, J.C.; Poirey, R.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: S67143
 A;Accession: S67149
 A;Molecule type: DNA
 A;Residues: 1-141 <JAU>
 A;Cross-references: UNIPROT:Q08687; EMBL:Z75160; NID:g1420570; PID:e252106; PID:g1420571
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: MIPS:YOR252w
 A;Cross-references: SGD:S0005778
 A;Map position: 15R

Query Match 23.5%; Score 23; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 115 NGTF 118

RESULT 33
 T49689
 hypothetical protein B23L21.20 [imported] - *Neurospora crassa*
 C;Species: *Neurospora crassa*
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
 C;Accession: T49689
 R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A;Reference number: Z25022
 A;Accession: T49689
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-143 <SCH>
 A;Cross-references: EMBL:ALJ356172; GSPDB:GN00116; NCSP:B23L21.20
 A;Experimental source: BAC clone B23L21; strain OR74A
 C;Genetics:

A;Gene: NCSP:B23L21.20
 A;Map position: 6
 A;Introns: 67/3; 80/2
 C;Superfamily: *Neurospora crassa* hypothetical protein B23L21.20

Query Match 23.5%; Score 23; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 71 NGTF 74

RESULT 34
 G91103
 hypothetical protein ECs3799 [imported] - *Escherichia coli* (strain O157:H7, substrain RN
 C;Species: *Escherichia coli*
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C;Accession: G91103
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: G91103
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-143 <HAY>
 A;Cross-references: UNIPROT:Q8XD02; GB:BA000007; PIDN:BA837222.1; PID:g13363271; GSPDB:GN
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECs3799

Query Match 23.5%; Score 23; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 67 NGTF 70

RESULT 35
 E71041
 hypothetical protein PH1621 - *Pyrococcus horikoshii*
 C;Species: *Pyrococcus horikoshii*
 C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
 C;Accession: E71041
 R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine,
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
 DNA Res. 5, 55-76, 1998
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic an
 A;Reference number: A71000; MUID:98344137; PMID:9679194
 A;Accession: E71041
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-143 <KAW>
 A;Cross-references: UNIPROT:O59287; GB:AP000006; NID:g3236133; PIDN:BAA30733.1; PID:g3236
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C;Genetics:
 A;Gene: PH1621

Query Match 23.5%; Score 23; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
 Db 77 NGTF 80

```
RESULT 36
B85949
hypothetical protein Z4267 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85949
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85949
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-143 <STO>
A:Cross-references: UNIPROT:Q8XD02; GB:AE005174; NID:g12517463; PIDN:AAG58054.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z4267

Query Match      23.5%; Score 23; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      67 NGTF 70
      ||||
      ||||

RESULT 37
T33494
hypothetical protein F40H3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33494
R:Gattung, S.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid F40H3.
A:Reference number: Z21358
A:Accession: T33494
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <GT>
A:Cross-references: UNIPROT:Q97ZK4; EMBL:AF098987; PIDN:AAC67428.1; GSPDB:GN000020; CESP:
A:Experimental source: strain Bristol N2; clone F40H3
C:Genetics:
A:Gene: CESP:F40H3.3
A:Map position: 2
A:Introns: 39/3; 72/1; 111/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F40H3.3

Query Match      23.5%; Score 23; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      65 NGTF 68
      ||||
      ||||

RESULT 38
S72492
probable proteinase inhibitor precursor - tomato
N:Alternate names: A72 protein
C:Species: Lycopersicon esculentum (tomato)
C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C:Accession: S72492
R:Brandstaedter, J.; Rossbach, C.; Theres, K.
Mol. Gen. Genet. 252, 146-154, 1996
A>Title: Expression of genes for a defensin and a proteinase inhibitor in specific areas
A:Reference number: S72491; MUID:96397493; PMID:8804387
A:Accession: S72492
A:Status: not compared with conceptual translation
```

```
A:Molecule type: mRNA
A:Residues: 1-146 <BRA>
A:Cross-references: UNIPROT:Q7W222
A:Experimental source: shoot, cv. Moneymaker
C:Superfamily: potato proteinase inhibitor PTI
C:Keywords: serine proteinase inhibitor
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-146/Product: probable proteinase inhibitor #status predicted <MAT>
F:33/Inhibitory site: Arg (trypsin) #status predicted
F:94/Inhibitory site: Arg (trypsin) #status predicted

Query Match      23.5%; Score 23; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      72 NGTF 75
      ||||
      ||||

RESULT 39
A69883
conserved hypothetical protein ylyA - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69883
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69883
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-147 <KUN>
A:Cross-references: UNIPROT:Q45478; GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13418.
A:Experimental source: strain 168
C:Genetics:
A:Gene: ylyA

Query Match      23.5%; Score 23; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      33 NGTF 36
Db      90 NGTF 93
      ||||
      ||||

RESULT 40
F69744
hypothetical protein ybbK - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: F69744
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
```

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69744
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <KUN>
A;Cross-references: UNIPROT:Q45584; GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11948.
A;Experimental source: strain 168
C;Genetics:
A;Gene: ybbK

Query Match 23.5%; Score 23; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 NGTF 36
Db 116 NGTF 119

Search completed: February 24, 2005, 15:24:26
Job time : 27.8537 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:04:15 ; Search time 104.715 Seconds
(without alignments)
273.851 Million cell updates/sec

Title: US-09-936-956-1
Perfect score: 98
Sequence: 1 XGXXXXXXXXXXXXXXXXX.....XXXXXXXXXXXXXXXXXXXXXG 56

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	30.6	204	2 Q9MBK0	Q9mbk0 bacterioph
2	29	29.6	822	2 Q73UA2	Q73ua2 mycobacteri
3	28	28.6	863	2 Q7DB14	Q7db14 mycobacteri
4	28	28.6	865	2 Q65934	Q65934 mycobacteri
5	28	28.6	865	2 Q7TZN0	Q7tzn0 mycobacteri
6	28	28.6	873	2 Q73UA3	Q73ua3 mycobacteri
7	27	27.6	854	2 Q73Z24	Q73z24 mycobacteri
8	26	26.5	132	2 Q8R8M6	Q8r8m6 thermoanaer
9	26	26.5	246	2 Q8DJ18	Q8dj18 synochococ
10	26	26.5	405	2 Q6AK27	Q6ak27 desulfotale
11	26	26.5	598	2 Q87PC3	Q87pc3 vibrio para
12	26	26.5	1543	2 Q71SN5	Q71sn5 helicobacte
13	25	25.5	220	2 Q86667	Q86667 streptomyc
14	25	25.5	260	2 Q6VRY4	Q6vry4 neisseria m
15	25	25.5	359	2 Q9RMN1	Q9rmn1 caulobacter
16	25	25.5	1025	1 SLAP_CAUCR	P35828 caulobacter
17	25	25.5	1105	2 Q703G3	Q703g3 pichia fari
18	24	24.5	65	2 Q62GC5	Q62gc5 burkholderi
19	24	24.5	65	2 Q63Q72	Q63q72 burkholderi
20	24	24.5	434	2 Q9KRW4	Q9khw4 salmonella
21	24	24.5	435	2 Q51801	Q51801 salmonella
22	24	24.5	455	2 Q6X909	Q6x909,salmonella
23	24	24.5	682	2 Q9EXB5	Q9exb5 neisseria m
24	24	24.5	684	2 Q3TPM0	Q9jpm0 neisseria m
25	24	24.5	686	2 Q9EXB6	Q9exb6 neisseria m
26	24	24.5	698	1 TBPB_NEIMA	Q88937 neisseria m
27	24	24.5	699	2 Q9EXC4	Q9exc4 neisseria m
28	24	24.5	1464	2 Q8V7J9	Q8v7j9 chlorella v
29	23	23.5	22	2 Q57226	Q57226 chlamydia t
30	23	23.5	22	2 Q79DL8	Q79dl8 chlamydia t
31	23	23.5	22	2 Q9R8K8	Q9r8k8 chlamydia t

32	23	23.5	22	2 Q9R8L3	Q9r8l3 chlamydia t
33	23	23.5	22	2 Q9R8L6	Q9r8l6 chlamydia t
34	23	23.5	22	2 Q9R8L9	Q9r8l9 chlamydia t
35	23	23.5	22	2 Q9R8M9	Q9r8m9 chlamydia t
36	23	23.5	22	2 Q9R8N3	Q9r8n3 chlamydia t
37	23	23.5	31	2 Q991P2	Q991p2 human immun
38	23	23.5	31	2 Q991P3	Q991p3 human immun
39	23	23.5	31	2 Q991P4	Q991p4 human immun
40	23	23.5	42	2 Q70UI4	Q70ui4 uncultured
41	23	23.5	42	2 Q710F3	Q710f3 uncultured
42	23	23.5	42	2 Q9JUA3	Q9juu3 neisseria m
43	23	23.5	44	2 Q7RDM6	Q7rdm6 plasmodium
44	23	23.5	48	2 Q8SAS5	Q8sas5 pinus sylve
45	23	23.5	49	2 Q8JIA6	Q8jia6 neotrophodiu

ALIGNMENTS

RESULT 1

ID	Q9MBK0	PRELIMINARY;	PRT;	204 AA.
AC	Q9MBK0;			
DT	01-OCT-2000 (Tremblrel. 15, Created)			
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Putative probed protease.			
GN	Name=ORF5;			
OS	Bacteriophage BFK20.			
OC	Viruses.			
OX	NCBI_TaxID=28358;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95185419; PubMed=7879712;			
RA	Kotides M., Ugorkakova J., Baloghova E., Bukovska G., Timko J.;			
RT	"Characterization and sequence analysis of the F2 promoter from corynephage BFK20.";			
RL	Acta Virol. 38:223-228(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92381485; PubMed=1512569;			
RA	Kotides M., Barak I., Sisova M., Baloghova E., Ugorkakova J.;			
RT	"Characterization of bacteriophage BFK20 from Brevibacterium flavum.";			
RL	J. Gen. Microbiol. 138:1387-1391(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Raslik I., Bukovska G., Godany A., Timko J.;			
RT	"Analysis of the nucleotide sequence of the genome of Brevibacterium flavum CC M 251 bacteriophage BFK20.";			
RL	Chem. Zvesti 52:266-267(1998).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Klucar L.;			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AJ278322; CAB93910.1; -.			
DR	MEROPS; U35.001; -.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	InterPro; IPR006433; Peptidase U35.			
DR	Pfam; PF04586; Peptidase U35; I.			
DR	TIGRFAMs; TIGR01543; proheadase_HK97; 1.			
KW	Protease.			
SQ	SEQUENCE. 204 AA; 22120 MW; 7B18F14718EF66E9 CRC64;			

Query Match 30.6%; Score 30; DB 2; Length 204;
Best Local Similarity 29.2%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFTAYASVFGNKSVDVVKSG 56
|||
Db 20 NGTFTAYASVFGNKSVDVVKSG 43

RESULT 2

```

Q73UA2
ID Q73UA2 PRELIMINARY; PRT; 822 AA.
AC Q73UA2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3466;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE017239; AAS06016.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR00253; FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00240; FHA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome; Hypothetical protein.
SQ SEQUENCE 822 AA; 87608 MW; 30747ED500ADC13F CRC64;

Query Match 29.6%; Score 29; DB 2; Length 822;
Best Local Similarity 29.2%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGDXXXXXXG 56
Db 232 NGTFVNGVRVCGAALSDGDVVITG 255

RESULT 3
ID Q7D814 PRELIMINARY; PRT; 863 AA.
AC Q7D814;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ABC transporter, ATP-binding protein.
GN OrderedLocusNames=MT1789;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and

```

```

RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE000516; AAK46062.1; -.
DR TIGR; MT1789; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR00253; FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00240; FHA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding.
SQ SEQUENCE 863 AA; 91924 MW; 48CB63FDC73C2DA3 CRC64;

Query Match 28.6%; Score 28; DB 2; Length 863;
Best Local Similarity 29.2%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGDXXXXXXG 56
Db 268 NGTFVNGARVDAALLHDGDVVITG 291

RESULT 4
ID O65934 PRELIMINARY; PRT; 865 AA.
AC O65934;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE CONSERVED TRANSMEMBRANE ATP-BINDING PROTEIN ABC
DE TRANSPORTER.
GN OrderedLocusNames=Rv1747;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; BX842577; CAB09333.1; -.
DR PIR; D70986; D70986.
DR Tuberculist; Rv1747; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.

```

```
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00240; FHA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 865 AA; 92152 MW; 4987186C471B47E8 CRC64;

Query Match 28.6%; Score 28; DB 2; Length 865;
Best Local Similarity 29.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGDXXXXXX 56
Db 270 NGTFVNGARVDALLHGDGVVTIG 293

RESULT 5
Q7TZNO PRELIMINARY; PRT; 865 AA.
ID Q7TZNO
AC Q7TZNO
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE CONSERVED TRANSMEMBRANE ATP-BINDING PROTEIN ABC
DE TRANSPORTER.
GN OrderedLocuNames=Mbl1776;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Biglimer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Ackin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; BX248340; CAD94478.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0042626; F: ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000253; FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC transporter; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome; Transmembrane.
SQ SEQUENCE 865 AA; 92152 MW; 4987186C471B47E8 CRC64;

Query Match 28.6%; Score 28; DB 2; Length 865;
Best Local Similarity 29.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGDXXXXXX 56
Db 270 NGTFVNGARVDALLHGDGVVTIG 293

RESULT 6
Q73UA3 PRELIMINARY; PRT; 873 AA.
ID Q73UA3
AC Q73UA3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=MAP3465;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Anonsin A., Alt D., Kapur V.;
RL Submitted (SEPP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE017239; AAS06015.1; -.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0042626; F: ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0001666; F: nucleotide binding; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR003593; AAA_ApPase.
DR InterPro; IPR003439; ABC_TRANSPORTER.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00240; FHA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome; Hypothetical protein.
SQ SEQUENCE 873 AA; 93301 MW; 280DF91EA9B263F CRC64;

Query Match 28.6%; Score 28; DB 2; Length 873;
Best Local Similarity 29.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGDXXXXXX 56
Db 279 NGTFVNGVRVGSVLTGDDVVTIG 302

RESULT 7
Q73Z24 PRELIMINARY; PRT; 854 AA.
ID Q73Z24
AC Q73Z24
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=MAP1457;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Anonsin A., Alt D., Kapur V.;
RL Submitted (SEPP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AE017232; AAS03774.1; -.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0042626; F: ATPase activity; coupled to transmembrane m. . .; IEA.
```

```

DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR000253; FHA__SMAD_FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00498; FHA; 2.
DR ProDom: PD000006; ABC_transporter; 1.
DR SMART: SM00240; FHA; 1.
DR SMART: SM00240; FHA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 854 AA; 89806 MW; 4F31B533AE308BAD CRC64;

Query Match 27.6%; Score 27; DB 2; Length 854;
Best Local Similarity 29.2%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXXXXG 56
Db 280 NGTFVNGVRVDSAVLRGDVVITG 303

RESULT 8
Q8R8M6 PRELIMINARY; PRT; 132 AA.
AC Q8R8M6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FHA-domain-containing proteins.
GN OrderedLocusNames=TTE1969;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=MB4;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Xu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013148; AAM25148.1; -.
DR HSSP; P46013; IR21.
DR InterPro: IPR00253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 132 AA; 15400 MW; 90936DC65A8B1158 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 132;
Best Local Similarity 29.2%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXXXXG 56
Db 101 NGTFVNGVRVKAIRKNGDIITLG 124

RESULT 9
Q8DJ18 PRELIMINARY; PRT; 246 AA.
AC Q8DJ18;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

10-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Tlr1235 protein.
OrderedLocusNames=tlr1235;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Nakaraki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AF005373; BAC08787.1; -.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 246 AA; 26355 MW; EAE4A94A64AD75C0 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 246;
Best Local Similarity 29.2%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXXXXG 56
Db 80 NGTFVNGKRISSELQHGVDVILFG 103

RESULT 10
Q6AK27 PRELIMINARY; PRT; 405 AA.
AC Q6AK27;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DP2570;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=LSv54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Anann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG37299.1; -.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 3.
DR SMART; SM00240; FHA; 3.
DR PROSITE; PS50006; FHA_DOMAIN; 3.
KW Complete proteome.
SQ SEQUENCE 405 AA; 43085 MW; 6BB6769AE076087E CRC64;

Query Match 26.5%; Score 26; DB 2; Length 405;
Best Local Similarity 31.8%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXGDXXXXXXG 51

```



```
Db 187 NGTFVAGKKTFAELRSGD 205
|||||
PRELIMINARY; PRT; 598 AA.

RESULT 11
Q87PC3
ID Q87PC3 PRELIMINARY; PRT; 598 AA.
AC Q87PC3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein VP1594.
OS OrderedLocusNames=VP1594;
GN Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005078; BAC59857.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 598 AA; 64664 MW; 3C41591D8AE218A6 CRC64;

Query Match 26.5%; Score 26; DB 2; Length 598;
Best Local Similarity 31.6%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGXD 51
|||||
PRELIMINARY; PRT; 1543 AA.

RESULT 12
Q71SN5
ID Q71SN5 PRELIMINARY; PRT; 1543 AA.
AC Q71SN5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE P167.
OS Helicobacter bilis.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=37372;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21981893; PubMed=11986271;
RA Feng S., Hodzic E., Kendall L.V., Smith A., Freet K., Barthold S.W.;
RT "Cloning and expression of a Helicobacter bilis immunoreactive
RT protein.";
RL Clin. Diagn. Lab. Immunol. 9:627-632(2002).
DR EMBL; AF288477; AA014336.1; -.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR005546; Auto.transpbeta.
DR Pfam; PF03797; Put. Vacuo cyctox.
DR Pfam; PF03077; Autotransporter; 1.
DR Pfam; PF03077; VacA2; 1.
SQ SEQUENCE 1543 AA; 167034 MW; 6CADA242PFF4297E CRC64;

Query Match 26.5%; Score 26; DB 2; Length 1543;
Best Local Similarity 25.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGXDXXXG 56
|||||
PRELIMINARY; PRT; 260 AA.

RESULT 14
Q6VRY4
ID Q6VRY4 PRELIMINARY; PRT; 260 AA.
AC Q6VRY4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Lipoprotein (Fragment).
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC-1573;
RX PubMed=15039331; DOI=10.1128/IAI.72.4.2088-2100.2004;
RA Fletcher L.D., Bernfield L., Barniak V., Farley J.E., Howell A.,
RA Knäuf M., Ooi P., Smith R.P., Weise P., Wetherell M., Xie X.,
RA Zagursky R., Zhang Y., Zlotnick G.W.;
RT "Vaccine potential of the Neisseria meningitidis 2086 lipoprotein.";
```

```
Db 95 NGTFNAKEVYITGLSGNAARTG 118
|||||
PRELIMINARY; PRT; 220 AA.

RESULT 13
O86667
ID O86667 PRELIMINARY; PRT; 220 AA.
AC O86667;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE Putative secreted protein.
GN ORFNames=SC4A2.11c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21964410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939127; CAA20163.1; -.
DR PIR; T34986; T34986.
DR HSSP; Q50769; INYO.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR00782; BIGH3 FAS1.
DR Pfam; PF02469; Fasciclin; 1.
DR SMART; SM00554; FAS1; 1.
DR PROSITE; PS0213; FAS1; 1.
DR Complete proteome.
SQ SEQUENCE 220 AA; 22362 MW; 78C08D7932B12C32 CRC64;

Query Match 25.5%; Score 25; DB 2; Length 220;
Best Local Similarity 26.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 33 NGTFPXXXXXXXGXD 51
|||||
PRELIMINARY; PRT; 260 AA.

Db 158 HGFTTLEGSTVTGSGGD 176
|||||
PRELIMINARY; PRT; 260 AA.

RESULT 14
Q6VRY4
ID Q6VRY4 PRELIMINARY; PRT; 260 AA.
AC Q6VRY4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Lipoprotein (Fragment).
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC-1573;
RX PubMed=15039331; DOI=10.1128/IAI.72.4.2088-2100.2004;
RA Fletcher L.D., Bernfield L., Barniak V., Farley J.E., Howell A.,
RA Knäuf M., Ooi P., Smith R.P., Weise P., Wetherell M., Xie X.,
RA Zagursky R., Zhang Y., Zlotnick G.W.;
RT "Vaccine potential of the Neisseria meningitidis 2086 lipoprotein.";
```

```

RL Infect. Immun. 72:2088-2100(2004).
DR EMBL; AY330406; AAR84481.1; -.
KW Lipoprotein.
FT CHAIN 1 1 260 lipoprotein.
SQ SEQUENCE 260 AA; 8CC49D53851B635F CRC64;

Query Match 25.5%; Score 25; DB 2; Length 260;
Best Local Similarity 25.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
Db 48 NGTLTLAQGAETGNGDSINTG 71

RESULT 15
Q9RWN1 PRELIMINARY; PRT; 359 AA.
AC Q9RWN1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Truncated S-layer protein.
GN Name=rsaA;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JS4000;
RX MEDLINE=21972786; PubMed=11976300;
RX DOI=10.1128/JB.184.10.2709-2718.2002;
RA Umelo-Njaka E., Bingle W.H., Borchani F., Le K.D., Awram P., Blake T.,
RA Nomellini J.F., Smit J.;
RT "Caulobacter crescentus synthesizes an S-layer-editing metalloprotease
RT possessing a domain sharing sequence similarity with its
RT paracrystalline S-layer protein.";
RL J. Bacteriol. 184:2709-2718(2002).
DR EMBL; AF193064; AAF07961.1; -.
SQ SEQUENCE 359 AA; 35434 MW; 62CA79A902F4B961 CRC64;

Query Match 25.5%; Score 25; DB 2; Length 359;
Best Local Similarity 25.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
Db 267 NDTPVAGEVAGATLTGVDTLGG 290

RESULT 16
SLAP CAUCR
ID SLAP CAUCR STANDARD; PRT; 1025 AA.
AC P35828; Q46015; Q9RF12;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE S-layer protein (Paracrystalline surface layer protein).
GN Name=rsaA; OrderedLocusNames=CC1007;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21; 646-660 AND 1021-1025.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=93007489; PubMed=1393820;
RA Gilchrist A., Fisher J.A., Smit J.K.;
RT "Nucleotide sequence analysis of the gene encoding the Caulobacter
RT crescentus paracrystalline surface layer protein.";
RL Can. J. Microbiol. 38:193-202(1992).

```

```

RN REVISIONS TO 376; 636 AND 842-843.
RP Awram P.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=JS3001;
RA Bingle W.H., Awram P.A., Nomellini J.F., Smit J.K.;
RT "The secretion signal of C. crescentus S-layer protein is located in
RT the C-terminal 82 amino acids of the molecule.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Niernan W.C., Feldblyum T.V., Iaub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Emolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
[5]
RP SEQUENCE OF 1-313 FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=89008089; PubMed=3049545;
RA Fisher J.A., Smit J.K., Agabian N.;
RT "Transcriptional analysis of the major surface array gene of
RT Caulobacter crescentus.";
RL J. Bacteriol. 170:4706-4713(1988).
[6]
RP CHARACTERIZATION.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=98292737; PubMed=9620954;
RA Awram P., Smit J.K.;
RT "The Caulobacter crescentus paracrystalline S-layer protein is
RT secreted by an ABC transporter (type I) secretion apparatus.";
RL J. Bacteriol. 180:3062-3069(1998).
CC -1- FUNCTION: The S-layer is a paracrystalline mono-layered assembly
CC of proteins which coats the surface of bacteria. Probably acts as
CC a physical barrier to parasites and lytic enzymes.
CC -1- SUBCELLULAR LOCATION: Cell wall. This bacterium is covered by a S-
CC layer with hexagonal symmetry. Secreted by an ABC transporter
CC (type I) secretion apparatus.
CC -1- MISCELLANEOUS: The calcium binding of this protein may mediate
CC surface attachment by ionic bridging between the S-layer and a
CC specific membrane-associated oligosaccharide.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF062345; AAC38665.2; -.
CC EMBL; AF193063; AAF19365.1; -.
CC EMBL; AE005779; AAK22991.1; ALT_INIT.
CC FIR; A48995; A48995.
CC HSP; P23629; ISR1.
CC TIGR; CC1007; -.
CC InterPro; IPR001343; Hemlysn_Ca_bind.
CC InterPro; IPR011049; Serralyen_like_C.
CC Pfam; PF00353; HemolysinCabin; 3.
CC PRINTS; PR00313; CABNDNGRPT.
CC Calcium-binding; Cell wall; Complete proteome;
CC Direct protein sequencing; S-layer.
CC INIT MET 0
CC SEQUENCE 1025 AA; 98001 MW; AD7A326E1363DB8AC CRC64;

```

```
Query Match      25.5%; Score 25; DB 1; Length 1025;
Best Local Similarity 25.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFFXXXXXXXGDXXXG 56
Db 266 NDTFVAGVAGAAATLTGDTLSGG 289

RESULT 17
Q703G3 PRELIMINARY; PRT; 1105 AA.
ID Q703G3;
AC Q703G3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE K, P-type ATPase.
GN Name=acul;
OS Pichia farinosa (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4920;
RN [1]
RP SEQUENCE FROM N.A.
RA Benito B., Garcadeblas B., Schriber P., Rodriguez-Navarro A.;
RA Submitted (JAN-2004) to the ENBL/GenBank/DBJ databases.
RL EMBL; AJ622831; CAF22247.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehalo like hydro.
DR InterPro; IPR008250; E1-E2 ATPase reg.
DR Pfam; PF00689; Cation_ATPase_N; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PROSITE; PS00121; NAKATPASE.
DR PROSITE; PS00154; ATPASE E1 E2; UNKNOWN 1.
SQ SEQUENCE 1105 AA; 122926 MW; 4878904DF7ECF50A CRC64;

Query Match      25.5%; Score 25; DB 2; Length 1105;
Best Local Similarity 26.1%; Pred. No. 2.4e+03;
Matches 6; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 34 GTFFXXXXXXXGDXXXG 56
Db 309 GTFCSTGSGTAIVSTGDNITFG 331

RESULT 18
Q62GC5 PRELIMINARY; PRT; 65 AA.
ID Q62GC5;
AC Q62GC5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Thiamine biosynthesis protein ThisS, putative.
GN ORFNames=BMA2729;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Nieman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Winn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammad Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU48294.1; -.
SQ SEQUENCE 65 AA; 6572 MW; 5364B131908DA8FE CRC64;

Query Match      24.5%; Score 24; DB 2; Length 65;
Best Local Similarity 26.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 NGTFFXXXXXXXGDX 51
Db 36 NGAFVARTQHARALAAAGD 54

RESULT 19
Q63Q72 PRELIMINARY; PRT; 65 AA.
ID Q63Q72;
AC Q63Q72;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Thiamine biosynthesis protein This.
GN Name=This; ORFNames=BPSL3153;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D., K.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH37163.1; -.
SQ SEQUENCE 65 AA; 6572 MW; 5364B131908DA8FE CRC64;

Query Match      24.5%; Score 24; DB 2; Length 65;
Best Local Similarity 26.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 33 NGTFFXXXXXXXGDX 51
Db 36 NGAFVARTQHARALAAAGD 54

RESULT 20
Q9KHW4 PRELIMINARY; PRT; 434 AA.
ID Q9KHW4;
AC Q9KHW4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PilV2.
```

```

GN Name=pilV2;
OS Salmonella dublin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Salmonella.
NCBI_TaxID=99360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serovar Dublin;
RX MEDLINE=22943013; PubMed=14580391; DOI=10.1016/j.micpath.2003.08.001;
RA Morris C., Tam C.K.P., Wallis T.S., Jones P.W., Hackett J.;
RT "Salmonella enterica serovar Dublin strains which are Vi antigen-
RT positive use type IVB pilli for bacterial self-association and human
RT intestinal cell entry.";
RL Microb. Pathog. 35:279-284 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=serovar Dublin;
RA Morris C.J., Hackett J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247502; AAF81218.1; -.
DR InterPro; IPR007001; Shufflon_N.
DR Pfam; PF04917; Shufflon_N; 1.
SQ SEQUENCE 434 AA; 46136 MW; 0273A4153BDD3518 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 434;
Best Local Similarity 25.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXGDXXXG 56
DB 238 NGNFGVSLVSNPITAGGDIRSTG 261

RESULT 21
OS1801
ID OS1801 PRELIMINARY; PRT; 435 AA.
AC OS1801; Q7ALS2; Q7C5H9;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE PilV (prepilin).
GN Name=pilV; OrderedLocustNames=STY4550, t4250;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Salmonella.
NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2;
RX MEDLINE=98087429; PubMed=9427557; DOI=10.1016/S0378-1119(97)00466-6;
RA Zhang X.L., Morris C., Hackett J.;
RT "Molecular cloning, nucleotide sequence, and function of a site-
RT specific recombinase encoded in the major 'pathogenicity island' of
RT Salmonella typhi.";
RL Gene 202:139-146 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2;
RA Zhang X.-L., Morris C., Hackett J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and Ct18.";
RL J. Bacteriol. 185:2330-2337 (2003).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Ct18;

```

```

RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jageis K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi Ct18.";
RL Nature 413:848-852 (2001).
DR EMBL; AF000001; AAC38069.1; -.
DR EMBL; AE016848; AAC071709.1; -.
DR EMBL; AL627282; CAD09328.1; -.
DR InterPro; IPR007001; Shufflon_N.
DR Pfam; PF04917; Shufflon_N; 1.
KW Complete proteome.
SQ SEQUENCE 435 AA; 46432 MW; C11F161143A93F79 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 435;
Best Local Similarity 25.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXGDXXXG 56
DB 238 NGNFGVSLVSNPITAGGDIRSTG 261

RESULT 22
Q6X909
ID Q6X909 PRELIMINARY; PRT; 455 AA.
AC Q6X909;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PilV2.
GN NamespilV2;
OS Salmonella enterica subsp. enterica serovar Paratyphi C.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Salmonella.
NCBI_TaxID=57046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CN13/87;
RX PubMed=14688076; DOI=10.1128/IAI.72.1.22-28.2004;
RA Tam C.K.P., Hackett J., Morris C.;
RT "Salmonella enterica serovar Paratyphi C carries an inactive
RT shufflon.";
RL Infect. Immun. 72:22-28 (2004).
DR EMBL; AY249242; AAP83041.1; -.
DR InterPro; IPR007001; Shufflon_N.
DR Pfam; PF04917; Shufflon_N; 1.
SQ SEQUENCE 455 AA; 48969 MW; 123098601188B5F6 CRC64;

Query Match 24.5%; Score 24; DB 2; Length 455;
Best Local Similarity 25.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 33 NGTFXXXXXXXXXXXXGDXXXG 56
DB 238 NGNFGVSLVSNPITAGGDIRSTG 261

RESULT 23
Q9EXB5
ID Q9EXB5 PRELIMINARY; PRT; 682 AA.
AC Q9EXB5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transferrin binding protein B subunit (fragment).
GN Name=tbpB;

```

OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=Z1500;
RX MEDLINE=21221108; PubMed=11287631; DOI=10.1073/pnas.0613860098;
RA Zhu P., van der Ende A., Falush D., Brieske N., Morelli G., Bodo L.,
RA Popovic T., Schuurman I.G.A., Adegbola R.A., Zurth K., Gagneux S.,
RA Platonov A.E., Riou J.Y., Caugant D.A., Nicolas P., Achtman M.;
RA "Fit genotypes and escape variants of subgroup III Neisseria
RT meningitidis during three pandemics of epidemic meningitis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5234-5239(2001).
DR EMBL; AJ276928; CAC21592.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004998; F:transferrin receptor activity; IEA.
DR InterPro: IPR001677; Transferrin_bind.
DR Pfam; PF01298; Lipoprotein_5; 1.
FT NON_TER
SQ SEQUENCE 684 AA; 73652 MW; D951B8868FCD7D6 CRC64;
Query Match 24.5%; Score 24; DB 2; Length 682;
Best Local Similarity 25.0%; Pred. No. 2.7e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Qy 33 NGTFXXXXXXXGDXXXG 56
Db 574 NGTLTADNRQAATFTVIGDIEGNG 597
RESULT 24
Q9JPMO
ID Q9JPMO PRELIMINARY; PRT; 684 AA.
AC Q9JPMO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transferrin binding protein B subunit (Fragment).
GN Name=cbpB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=Z3906;
RX MEDLINE=21221108; PubMed=11287631; DOI=10.1073/pnas.0613860098;
RA Zhu P., van der Ende A., Falush D., Brieske N., Morelli G., Bodo L.,
RA Popovic T., Schuurman I.G.A., Adegbola R.A., Zurth K., Gagneux S.,
RA Platonov A.E., Riou J.Y., Caugant D.A., Nicolas P., Achtman M.;
RA "Fit genotypes and escape variants of subgroup III Neisseria
RT meningitidis during three pandemics of epidemic meningitis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5234-5239(2001).
DR EMBL; AJ276909; CAB3048.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004998; F:transferrin receptor activity; IEA.

DR InterPro: IPR001677; Transferrin_bind.
DR Pfam; PF01298; Lipoprotein_5; 1.
FT NON_TER
SQ SEQUENCE 684 AA; 73610 MW; 13D6B69FBD5A425 CRC64;
Query Match 24.5%; Score 24; DB 2; Length 684;
Best Local Similarity 25.0%; Pred. No. 2.7e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Qy 33 NGTFXXXXXXXGDXXXG 56
Db 576 NGTLTADNRQAATFTVIGDIEGNG 599
RESULT 25
Q9EXB6
ID Q9EXB6 PRELIMINARY; PRT; 686 AA.
AC Q9EXB6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transferrin binding protein B subunit (fragment).
GN Name=cbpB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=Z5828;
RX MEDLINE=20305048; PubMed=10844690;
RA Linz B., Schenker M., Achtman M.;
RA "Frequent interspecific genetic exchange between commensal Neisseria
RT and Neisseria meningitidis.";
RL Mol. Microbiol. 36:1049-1058(2000).
RN
RN SEQUENCE FROM N.A.
RC STRAIN=Z5828;
RX MEDLINE=21221108; PubMed=11287631; DOI=10.1073/pnas.0613860098;
RA Zhu P., van der Ende A., Falush D., Brieske N., Morelli G., Bodo L.,
RA Popovic T., Schuurman I.G.A., Adegbola R.A., Zurth K., Gagneux S.,
RA Platonov A.E., Riou J.Y., Caugant D.A., Nicolas P., Achtman M.;
RA "Fit genotypes and escape variants of subgroup III Neisseria
RT meningitidis during three pandemics of epidemic meningitis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5234-5239(2001).
DR EMBL; AJ276927; CAC21591.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004998; F:transferrin receptor activity; IEA.
DR InterPro: IPR001677; Transferrin_bind.
DR Pfam; PF01298; Lipoprotein_5; 1.
FT NON_TER
SQ SEQUENCE 686 AA; 73774 MW; 5DC31FEB12AF1E32 CRC64;
Query Match 24.5%; Score 24; DB 2; Length 686;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Qy 33 NGTFXXXXXXXGDXXXG 56
Db 578 NGTLTADNRQAATFTVIGDIEGNG 601
RESULT 26
TEBP_NEIMA
ID TBPB_NEIMA STANDARD; PRT; 698 AA.
AC O68937;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Transferrin-binding protein 2 precursor (TBP-2).
GN Name=cbpB; Synonyms=tbp2; OrderedLocusNames=NMA2025;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

```
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RA Brieske N., Quentin-Millet M.-J., Schenker M., Schnibbe T.,
RA Achtman M.;
RT "Human antibody responses to A and C capsular polysaccharides, IgA1
RT protease and transferrin-binding protein complex stimulated by
RT infection with Neisseria meningitidis of subgroup IV-1 or ET-37
RT complex.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham K.D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
CC -!- FUNCTION: Acts as a transferrin receptor and is required for
CC transferrin utilization (By similarity).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF058689; AAC13725.1; -;
DR EMBL; AL162757; CAB85244.1; -;
DR PIR; D81832; D81832.
DR InterPro; IPR000437; Prok lipoprot S.
DR InterPro; IPR001677; Transferrin_bind.
DR Pfam; PF01298; Lipoprotein 5; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; FALSE_NEG.
KW Complete proteome; Lipoprotein; Outer membrane; Palmitate; Receptor;
KW Signal.
FT SIGNAL 1 20 By similarity.
FT CHAIN 21 698 Transferrin-binding protein 2.
FT LIPID 21 21 N-palmitoyl cysteine (Probable).
FT LIPID 21 21 S-diacylglycerol cysteine (Probable).
SQ SEQUENCE 698 AA; 75176 MW; 83FBB14DDF617B1F CRC64;

Query Match 24.5%; Score 24; DB 1; Length 698;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
Db 590 NGTLTADNRQAATFTIVGDIEGNG 613

RESULT 27
Q9EXC4 PRELIMINARY; PRT; 699 AA.
AC Q9EXC4
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transferrin binding protein B subunit (fragment).
GN Name:tbps;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
```

```
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z5642;
RX MEDLINE=20305048; PubMed=10844690;
RA Linz B., Schenker M., Achtman M.;
RT "Frequent interspecific genetic exchange between commensal Neisseriae
RT and Neisseria meningitidis.";
RL Mol. Microbiol. 36:1049-1058(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z5642;
RX MEDLINE=21221108; PubMed=11287631; DOI=10.1073/pnas.061386098;
RA Zhu P., van der Ende A., Falush D., Brieske N., Morelli G., Bodo L.,
RA Popovic T., Schuurman I.G.A., Adegbola R.A., Zurth K., Gagneux S.,
RA Platonov A.B., Riou J.Y., Caugant D.A., Nicolas P., Achtman M.;
RT "Fit genotypes and escape variants of subgroup III Neisseria
RT meningitidis during three pandemics of epidemic meningitis.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5234-5239(2001).
DR EMBL; AJ276919; CAC21583.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004998; F:transferrin receptor activity; IEA.
DR InterPro; IPR001677; Transferrin_bind.
DR Pfam; PF01298; Lipoprotein 5; 1.
FT NON_TER 1 1
SQ SEQUENCE 699 AA; 75373 MW; 60B6F8CECA208ADE CRC64;

Query Match 24.5%; Score 24; DB 2; Length 699;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 6; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXGDXXXG 56
Db 591 NGTLTADNRQAATFTIVGDIEGNG 614

RESULT 28
Q8V7J9 PRELIMINARY; PRT; 1464 AA.
AC Q8V7J9
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Vp260 like protein.
GN Name:ORF3;
OS Chlorella virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10507;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CVK2;
RX MEDLINE=22031233; PubMed=12033788; DOI=10.1006/viro.2002.1408;
RA Chuchird N., Nishida K., Kawasaki T., Fujie M., Usami S., Yamada T.;
RT "A variable region on the chlorovirus CVK2 genome contains five copies
RT of the gene for Vp260, a viral-surface glycoprotein.";
RL Virology 295:289-298(2002).
DR EMBL; AB063108; BAB83469.1; -;
DR Pfam; PF06598; Chlorovi_gp_rpt; 14.
SQ SEQUENCE 1464 AA; 149988 MW; ECCDA711437B6F5D CRC64;

Query Match 24.5%; Score 24; DB 2; Length 1464;
Best Local Similarity 27.8%; Pred. No. 5.7e+03;
Matches 5; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 33 NGTFXXXXXXXXXXXXXG 50
Db 1331 NGTFSNYGALTGTSTAG 1348

RESULT 29
Q57226 PRELIMINARY; PRT; 22 AA.
ID Q57226
```

```
AC Q57226;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95318541; PubMed=7797929;
RA Hayes L.J., Pecharatana S., Bailey R.L., Hampton T.J., Pickett M.A.,
RA Mabey D.C., Watt P.J., Ward M.E.;
RT "Extent and kinetics of genetic change in the omp1 gene of Chlamydia
RT trachomatis in two villages with endemic trachoma.";
RL J. Infect. Dis. 172:268-272(1995).
DR EMBL; S77984; AAB34697.1; -.
FT NON_TER 22
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 30
Q79DL8 PRELIMINARY; PRT; 22 AA.
AC Q79DL8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Omp1 (Fragment).
GN Name=omp1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94044912; PubMed=7901288;
RA Yang C.L., Maclean I., Brunham R.C.;
RT "DNA sequence polymorphism of the Chlamydia trachomatis omp1 gene.";
RL J. Infect. Dis. 168:1225-1230(1993).
DR EMBL; S66728; AAB28611.1; -.
FT NON_TER 22
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 31
Q9R8K8 PRELIMINARY; PRT; 22 AA.
AC Q9R8K8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 32
Q9R8L3 PRELIMINARY; PRT; 22 AA.
AC Q9R8L3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B23a;
RA Hsieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070276; AAC25246.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 33
Q9R8L6 PRELIMINARY; PRT; 22 AA.
AC Q9R8L6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B24b;
RA Hsieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070272; AAC25242.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B2f;
RA Hsieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070284; AAC25254.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 32
Q9R8L3 PRELIMINARY; PRT; 22 AA.
AC Q9R8L3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B23a;
RA Hsieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070276; AAC25246.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 33
Q9R8L6 PRELIMINARY; PRT; 22 AA.
AC Q9R8L6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B24b;
RA Hsieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070272; AAC25242.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 34
Q9R8L9 ID Q9R8L9 PRELIMINARY; PRT; 22 AA.
AC Q9R8L9 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2d;
RA Heieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070268; AAC25238.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 35
Q9R8M9 ID Q9R8M9 PRELIMINARY; PRT; 22 AA.
AC Q9R8M9 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B124a;
RA Heieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070248; AAC25218.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 36
Q9R8N3 ID Q9R8N3 PRELIMINARY; PRT; 22 AA.
AC Q9R8N3 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B12a;
RA Heieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070244; AAC25214.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13

RESULT 37
Q991P2 ID Q991P2 PRELIMINARY; PRT; 31 AA.
AC Q991P2 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein V4 region (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21109142; PubMed=11162839; DOI=10.1006/viro.2000.0753;
RA Ghezzi S., Menzo S., Brambilla A., Panina Bordignon P., Lorini A.L.,
RA Clementi M., Poli G., Vicenzi E.;
RT "Inhibition of R5X4 dualtropic HIV-1 primary isolates by single
RT chemokine co-receptor ligands.";
RL Virology 280:253-261(2001).
DR EMBL; AF333223; AAK21198.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 31
SQ SEQUENCE 31 AA; 3396 MW; 1FE12AA0F6C50B9A CRC64;

Query Match 23.5%; Score 23; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 24 NGTF 27

RESULT 38
Q991P3 ID Q991P3 PRELIMINARY; PRT; 31 AA.
AC Q991P3 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Major outer membrane protein (Fragment).
GN Name=omp-1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B12a;
RA Heieh Y.-H., Bobo L.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070244; AAC25214.1; -.
FT NON_TER 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2425 MW; B102482844167521 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 10 NGTF 13
```



```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein V4 region (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21109142; PubMed=11162839; DOI=10.1006/viro.2000.0753;
RA Ghezzi S., Menzo S., Brambilla A., Panina Bordignon P., Lorini A.L.,
RA Clementi M., Poli G., Vicenzi E.;
RT "Inhibition of RSX4 dualtropic HIV-1 primary isolates by single
RT chemokine co-receptor ligands.";
RL Virology 280:253-261(2001).
DR EMBL; AF333222; AA21197.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR GO; GO:0005198; F:structural molecule activity; IEA.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3395 MW; EF412AA0F6C50B97 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 24 NGTF 27

RESULT 39
Q991P4 PRELIMINARY; PRT; 31 AA.
AC Q991P4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein V4 region (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21109142; PubMed=11162839; DOI=10.1006/viro.2000.0753;
RA Ghezzi S., Menzo S., Brambilla A., Panina Bordignon P., Lorini A.L.,
RA Clementi M., Poli G., Vicenzi E.;
RT "Inhibition of RSX4 dualtropic HIV-1 primary isolates by single
RT chemokine co-receptor ligands.";
RL Virology 280:253-261(2001).
DR EMBL; AF333221; AA21196.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3383 MW; EF412AB106C50B97 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 33 NGTF 36
Db 24 NGTF 27

RESULT 40
Q70UI4 PRELIMINARY; PRT; 42 AA.
AC Q70UI4
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Laccase (EC 1.10.3.2) (Fragment).
GN Name=lac;
OS uncultured basidiomycete.
OC Eukaryota; Fungi; Basidiomycota; environmental samples.
OX NCBI_TaxID=175244;
RN [1]
RP SEQUENCE FROM N.A.
RA Luis P., Walther G., Kellner H., Martin F., Buscot F.;
RT "Diversity of laccase genes from basidiomycetes in a forest soil.";
RL Soil Biol. Biochem. 36:1025-1036(2004).
DR EMBL; AJ540279; CAD62542.1; -.
DR HSP; Q9Y780; 1A65.
DR GO; GO:0008471; F:laccase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR008972; Cupredoxin.
KW Oxidoreductase.
FT NON_TER 1 1
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 5033 MW; 4381115F42638C25 CRC64;

Query Match 23.5%; Score 23; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 NGTF 36
Db 35 NGTF 38

Search completed: February 24, 2005, 15:23:35
Job time : 113.715 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:03:10 ; Search time 140.537 Seconds
(without alignments)
184.386 Million cell updates/sec

Title: US-09-936-956-19

Perfect score: 336

Sequence: 1 WTRGNPACDYHLGNISRLS.....QKVEKNSQLLSQDEITVG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseqp16Dec04:*
2: Geneseqp1980s:*
3: Geneseqp1980s:*
4: Geneseqp2000s:*
5: Geneseqp2001s:*
6: Geneseqp2002s:*
7: Geneseqp2003as:*
8: Geneseqp2003bs:*
9: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	313.5	93.3	776	2	AAW73894
2	313.5	93.3	821	2	AAW26664
3	313.5	93.3	821	5	ABP54942
4	313.5	93.3	821	6	ABR53872
5	313.5	93.3	821	7	ABU61612
6	313.5	93.3	821	7	ADK64880
7	168	50.0	699	3	ABO3445
8	168	50.0	699	5	ABP73798
9	93	27.7	609	8	ADS44244
10	86.5	25.7	474	8	ADN19646
11	85.5	25.4	558	7	ABO68456
12	85	25.3	426	4	AAU15856
13	85	25.3	426	6	ABU54925
14	85	25.3	572	8	ABM83354
15	85	25.3	612	8	ABM83353
16	85	25.3	612	8	ABM83352
17	85	25.3	623	4	AAAB93182
18	85	25.3	623	5	ABB97233
19	85	25.3	652	4	AAAB93168
20	85	25.3	664	4	AAAB83843
21	85	25.3	664	4	AAAB20219
22	84	25.0	543	3	AAAY67968
23	81.5	24.3	288	4	AAAG89798
24	79.5	23.7	148	2	AAW32352
25	79.5	23.7	148	2	AAW32420

26	79.5	23.7	148	2	AAW64292	Myobacte
27	79.5	23.7	148	2	AAW81655	M. tuberc
28	79.5	23.7	148	2	AAAY38957	M. tuberc
29	79.5	23.7	148	2	AAAY39094	M. tuberc
30	79.5	23.7	161	2	AAAY72886	Myobacte
31	79.5	23.7	162	2	AAAY21903	Amino aci
32	79.5	23.7	162	2	AAAY50735	Myobacte
33	77	22.9	514	5	AAAM48994	Human Chk
34	77	22.9	514	8	ADL32563	Human Chk
35	77	22.9	542	4	AAAG68375	Human Chk
36	77	22.9	543	2	AAAY06204	Human che
37	77	22.9	543	2	AAAY05765	Novel hum
38	77	22.9	543	5	AAAM48996	Human Chk
39	77	22.9	543	5	AAAM48995	Human Chk
40	77	22.9	543	5	ABG30701	Human che
41	77	22.9	543	8	ADL32561	Human Chk
42	77	22.9	543	8	ADMT72213	Human TAS
43	77	22.9	543	8	ADO44010	Amino aci
44	77	22.9	543	8	ADOQ9232	Human CHE
45	77	22.9	586	8	ADN61455	Human KPP

ALIGNMENTS

RESULT 1

AAW73894
ID AAW73894 standard; protein; 776 AA.

XX AC AAW73894;

XX AC

DT 08-APR-1999 (first entry)

XX DE Yeast MEC2 protein sequence.

XX DE Cell cycle checkpoint gene; yeast; radiation resistance; detection;

XX DE G2/M checkpoint.

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

CC cell is inhibited at the restrictive temperature, comprising: (a)
 CC obtaining a human cDNA library comprising several human cDNA clones; (b)
 CC inserting the human cDNA clones individually into plasmid vectors
 CC containing a selectable marker gene; (c) transforming a culture of the
 CC yeast test cells with the plasmid vectors from the preceding step; (d)
 CC selecting for yeast test cells transformed with the selectable marker
 CC gene; (e) growing the selected transformants at the restrictive
 CC temperature and isolating a candidate transformant capable of growing at
 CC the restrictive temperature; and (f) identifying the human cDNA carried
 CC by the candidate transformant as a human checkpoint cDNA by sequencing
 CC the human cDNA carried by the candidate transformant and determining that
 CC the human cDNA is less than 50% homologous with both the first gene and
 CC the second gene. The nucleotide sequence of huRADcompB can be used to
 CC confer radiation resistance on a cell
 XX
 SQ Sequence 776 AA;

Query Match 93.3%; Score 313.5; DB 2; Length 776;
 Best Local Similarity 89.7%; Pred. No. 2.4e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTTLNGQKVEKNSNOLLS 59
 Db 66 WTFGRNPACDYHLGNISRLSNKHFOILLGDEGNLLNDISTNGTTLNGQKVEKNSNOLLS 125
 Qy 60 QGDEITVG 67
 Db 126 QGDEITVG 133

RESULT 2
 AAW26664
 ID AAW26664 standard; protein; 821 AA.
 AC AAW26664;
 XX
 XX 25-MAR-2003 (revised)
 DT 25-FEB-1998 (first entry)
 XX
 XX Yeast checkpoint control protein MEC2.
 XX
 XX MEC2; cell cycle; checkpoint gene; yeast; DNA damage; cancer; therapy.
 XX
 XX Saccharomyces cerevisiae.

Key Location/Qualifiers
 PH Protein 1..776
 FT /note= "in the Sequence Listing MEC2 terminates at
 FT residue 775"
 XX
 XX US5674996-A.
 XX
 XX 07-OCT-1997.
 XX
 XX 18-FEB-1994; 94US-00198446.
 XX
 XX 12-MAY-1992; 92US-00882051.
 PR 14-MAY-1992; 92US-00884426.
 PR 12-MAY-1993; 93WO-US004458.
 XX

(UYAR-) UNIV ARIZONA.
 PA (UNIW) UNIV WASHINGTON.
 PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 XX
 XX Groudine MT, Weinert TA, Plon SE, Hartwell LH;
 XX
 XX WPI; 1997-502392/46.
 DR N-PSDB; AAT91040.
 XX
 XX Nucleotide sequence capable of hybridising with huCDC34 - which is human
 PT checkpoint gene, useful to increase sensitivity of tumour cells to
 PT chemotherapeutic drugs or radiation.
 XX

Example 4; Col 85-90; 54pp; English.

PS This protein is encoded by novel yeast checkpoint control gene MEC2 (see
 XX AAT91040). Yeast checkpoint control genes RAD17, RAD24, MEC1, MEC2 and
 CC MEC3 (AAT91037-41, respectively) are responsible for recognising if the
 CC cell has suffered DNA damage in the form of radiation or chemical damage
 CC or if the cell has failed to complete DNA replication because of chemical
 CC inhibition or intrinsic error. Upon recognising damage or failure, the
 CC genes are responsible for inhibiting mitosis. The purpose of this
 CC checkpoint control is that it preserves the viability of the cell and the
 CC integrity of the genome by providing the cell time to repair these
 CC insults prior to undertaking mitosis. The genes are potentially useful in
 CC developing cancer chemotherapeutics, cancer chemoprevention agents, and
 CC environmental toxicology tests. They can be used to produce proteins (see
 CC AAW26661-65) that can then be screened for chemical agents that would
 CC interfere with checkpoint controls. Cloned genes can also be used to
 CC develop yeast strains in which these genes are deleted. Such yeast
 CC strains can then be used to find the homologous human genes (see AAT91034
 CC -36). (Updated on 25-MAR-2003 to correct PF field.)
 XX

SQ Sequence 821 AA;

Query Match 93.3%; Score 313.5; DB 2; Length 821;
 Best Local Similarity 89.7%; Pred. No. 2.6e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTTLNGQKVEKNSNOLLS 59
 Db 66 WTFGRNPACDYHLGNISRLSNKHFOILLGDEGNLLNDISTNGTTLNGQKVEKNSNOLLS 125
 Qy 60 QGDEITVG 67
 Db 126 QGDEITVG 133

RESULT 3
 ABP54942
 ID ABP54942 standard; peptide; 821 AA.
 XX
 AC ABP54942;

XX 13-JAN-2003 (first entry)
 DT
 XX Saccharomyces cerevisiae SPK1.
 DE
 XX SPK1; tyrosine threonine kinase; TTK; protein kinase; enzyme;
 KW mitotic checkpoint; colon cancer; breast cancer; tumour; cytostatic;
 KW gene therapy.
 XX
 XX Saccharomyces cerevisiae.
 OS
 XX WO200268444-A1.
 PN
 XX 06-SEP-2002.
 PD
 XX 21-FEB-2002; 2002WO-US005278.
 PF
 XX 21-FEB-2001; 2001US-0271254P.
 PR
 XX (CHIR) CHIRON CORP.
 PA
 XX Reinhard C, Jefferson AB, Chan VW;
 PI
 XX WPI; 2002-698650/75.
 DR N-PSDB; ABV73988.
 DR

XX Reducing growth of cancer cells comprises reducing Tyrosine Threonine
 PT Kinase (TTK) activity, useful in diagnosing and treating disorders with
 PT abnormal expression levels and activity of TTK, such as lung, colon,
 PT prostate and ovarian cancer.
 XX
 XX Disclosure; Page 77-78; 113pp; English.
 XX

CC The present sequence is the protein sequence of *Saccharomyces cerevisiae*
 CC SPK1, a protein related to human tyrosine threonine kinase (TTK, see
 CC ABP54938). TTK polynucleotides and polypeptides of the invention
 CC encompass polynucleotides and polypeptides having sequence similarity or
 CC sequence identity to human TTK and other genes and gene products related
 CC to TTK, such as SPK1. The invention is based on the finding that TTK is
 CC differentially expressed in various forms of cancer. It provides methods
 CC for the identification of cancerous cells, especially breast cancer and
 CC colon cancer cells, by detection of expression levels of TTK, as well as
 CC diagnostic, prognostic and therapeutic methods. These methods can be used
 CC as the basis of rational therapy. Assays for identifying molecules that
 CC modulate the activity of these genes in cancers, as well as methods of
 CC inhibiting tumour growth by inhibiting the activity of TTK are also
 CC provided

XX Sequence 821 AA;

Query Match 93.3%; Score 313.5; DB 5; Length 821;
 Best Local Similarity 89.7%; Pred. No. 2.6e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTCGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWLNGQVKVKNQNLLS 59
 |||||
 Db 66 WTCGRNPACDYHLGNISRLSNKHFOILLGEGDGNLLNDISTNGTWLNGQVKVKNQNLLS 125

Qy 60 QGDEITVG 67
 |||||
 Db 126 QGDEITVG 133

RESULT 4

ABR53872
 ID ABR53872 standard; protein; 821 AA.

XX AC ABR53872;

XX DT 20-JUN-2003 (first entry)

XX DE Protein sequence #SEQ ID 2609.

XX KW Multiprotein complex; eukaryote; drug target; diagnosis.

XX OS *Saccharomyces cerevisiae*.

XX FN EP1258494-A1.

XX PD 20-NOV-2002.

XX PF 20-DEC-2001; 2001EP-00130253.

XX PR 15-MAY-2001; 2001EP-00111774.

XX PA (CELL-) CELLZOME AG.

PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
 PI Marzioch M, Schultz JD, Superti-Furga GD;

XX WPI; 2003-250078/25.

XX DR N-PSDB; ACC61914.

PT New isolated protein complexes useful for diagnosing a disease or
 PT disorder, or as a target for an active agent of a pharmaceutical,
 PT preferably a drug target in the treatment or prevention of disease or
 PT disorder.

XX Disclosure; SEQ ID NO 2609; 17pp + Sequence Listing; English.

CC The invention relates to multiprotein complexes from eukaryotes. Proteins
 CC of the invention and DNA sequences encoding them are given in records
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
 CC obtainable by using a protein as a bait and isolating the set of proteins
 CC which is attached thereto from cells. Such protein complexes may comprise
 CC up to 30 distinct proteins. Protein complexes of the invention are useful

CC for diagnosing a disease or disorder, or as a target for an active agent
 CC of a pharmaceutical, preferably a drug target in the treatment or
 CC prevention of a disease or disorder. Note: The sequence data for this
 CC patent is not represented in the printed specification, but is based on
 CC sequence information supplied by the European Patent Office. The complete
 CC document is available on CD-ROM

SQ Sequence 821 AA;

Query Match 93.3%; Score 313.5; DB 6; Length 821;
 Best Local Similarity 89.7%; Pred. No. 2.6e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTCGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWLNGQVKVKNQNLLS 59
 |||||
 Db 66 WTCGRNPACDYHLGNISRLSNKHFOILLGEGDGNLLNDISTNGTWLNGQVKVKNQNLLS 125

Qy 60 QGDEITVG 67
 |||||
 Db 126 QGDEITVG 133

RESULT 5

ABU61612
 ID ABU61612 standard; protein; 821 AA.

XX AC ABU61612;

XX DT 11-AUG-2003 (first entry)

XX DE Yeast SPK1 protein.

XX KW Yeast; tyrosine threonine kinase; TTK; cancer; cytostatic;
 KW mitotic checkpoint gene; SPK1.

XX OS *Saccharomyces cerevisiae*.

XX FN US2003045491-A1.

XX PD 06-MAR-2003.

XX PF 21-FEB-2002; 2002US-00081119.

XX PR 23-FEB-2001; 2001US-0289813P.

XX PA (REIN/) REINHARD C.

XX PA (JEFF/) JEFFERSON A B.

XX PA (CHAN/) CHAN V W.

XX PI Reinhard C, Jefferson AB, Chan VW;

XX DR WPI; 2003-456566/43.

XX DR N-PSDB; ACA62264.

XX Detecting cancer in a subject, by comparing expression levels of tyrosine
 PT threonine kinase polypeptide or polynucleotide in a subject cell and a
 PT normal cell, where an increase in the expression level in the test cell
 PT is indicative of cancer.

XX Disclosure; Page 30-32; 79pp; English.

XX The invention relates to detecting cancer (other than ovarian cancer) in
 CC a subject, comprising comparing the expression levels of tyrosine
 CC threonine kinase (TTK, a mitotic checkpoint gene) polypeptide or
 CC polynucleotide in a test cell obtained from the subject and in a normal
 CC non-cancer cell, where an increase in the expression level of TTK protein
 CC or nucleic acid in the test cell compared to that in the normal cell,
 CC indicates the presence of cancer other than ovarian cancer. Also included
 CC are reducing growth of a cancerous cell (by contacting a cancerous cell
 CC with an amount of an agent effective to reduce TTK polypeptide activity
 CC in the cell), an assay for identifying a candidate agent that reduces
 CC growth of a cancerous cell (comprising: (i) detecting the activity of a
 CC TTK polypeptide in the presence of a candidate agent; and (ii) comparing

CC the activity of TTK polypeptide in the presence of a candidate agent
 CC relative to TTK polypeptide activity in the absence of the candidate
 CC agent), identifying an agent that reduces TTK activity (comprising: (i)
 CC contacting a cancerous cell displaying elevated expression of a TTK-
 CC encoding polynucleotide with a candidate agent; and (ii) determining the
 CC effect of the candidate agent on TTK polypeptide activity) and assessing
 CC the prognosis of a cancerous disease other than ovarian cancer in a
 CC subject (comprising: (i) detecting expression of TTK -encoding
 CC polynucleotide in a test cancer cell of a subject; and (ii) comparing a
 CC level of expression of TTK-encoding polynucleotide in the test cancer
 CC cell with a level of expression of the polynucleotide in a control non-
 CC cancer cell, where the level of expression of TTK in the test cancer cell
 CC relative to the level of expression in the control non-cancer cell is
 CC indicative of the prognosis of the cancerous disease). The methods are
 CC useful for detecting cancer (other than ovarian cancer) in a subject.
 CC reducing growth of cancerous cells, identifying a candidate agent that
 CC reduces growth of a cancerous cell, identifying an agent that reduces TTK
 CC activity and assessing the prognosis of a cancerous disease other than
 CC ovarian cancer. The methods are also useful for determining the ability
 CC of a subject to respond to a particular therapy e.g. as a basis of
 CC rational therapy. The present sequence represents a closely related
 CC protein to human TTK, in this case yeast SPK1 (not defined)
 XX
 SQ Sequence 821 AA;

Query Match 93.3%; Score 313.5; DB 7; Length 821;
 Best Local Similarity 89.7%; Pred. No. 2.6e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
 QY 1 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNLGQKVERKNSQLLS 59
 DB 66 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNLGQKVERKNSQLLS 125
 QY 60 QGDEITVG 67
 DB 126 QGDEITVG 133

RESULT 6
 ADK64880
 ID ADK64880 standard; protein; 821 AA.
 XX
 AC ADK64880;

DT 06-MAY-2004 (first entry)
 DE Disease treating protein complex-derived protein #1576.
 XX protein complex; drug target; diagnosis.

XX Unidentified.

XX EPI338608-A2.

XX 27-AUG-2003.

XX 20-DEC-2002; 2002EP-00102902.

XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELLZOME AG.

XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
 PI Michon A, Leutwein C, Rick J;

XX WPI; 2003-639460/61.

DR N-PSDB; ADK64881.

XX New proteins and protein complexes from eukaryotes, useful as targets in
 PT drug screening, or in diagnosing or screening for the presence of a
 PT disease or disorder, or a predisposition for developing a disease or
 PT disorder in a subject.

XX Disclosure; SEQ ID NO 3151; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a
 CC second protein, or its derivative, fragment, homologue or variant. The
 CC proteins are selected from given protein complexes, which are not defined
 CC in the specification. The variants are encoded by nucleic acids that
 CC hybridize to the nucleic acids encoding the proteins under low stringency
 CC conditions. The protein complexes are useful as targets for an active
 CC agent of a pharmaceutical. These protein complexes are particularly
 CC useful as drugs targets for the treatment or preventing of a disease or
 CC disorder. The complexes and methods above are useful in diagnosing or
 CC screening for the presence of a disease or disorder or a predisposition
 CC for developing a disease or disorder in a subject. These are also useful
 CC in screening for a drug for treatment or prevention of a disease or
 CC disorder. The molecule that modulates the amount, activity or protein
 CC components of the complex is useful for the manufacture of a medicament
 CC for the treatment or prevention of a disease or disorder. This sequence
 CC corresponds to a protein of the invention. (Note: the sequence data for
 CC this patent did not form part of the printed specification but was
 CC obtained from the EPO in electronic format).

XX
 SQ Sequence 821 AA;

Query Match 93.3%; Score 313.5; DB 7; Length 821;
 Best Local Similarity 89.7%; Pred. No. 2.6e-33;
 Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNLGQKVERKNSQLLS 59
 DB 66 WTFGRNPACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNLGQKVERKNSQLLS 125
 QY 60 QGDEITVG 67
 DB 126 QGDEITVG 133

RESULT 7
 AAB03445
 ID AAB03445 standard; protein; 699 AA.
 XX
 AC AAB03445;

DT 03-JAN-2001 (first entry)

DE Candida albicans essential growth protein #3.

XX Growth inhibition; survival; pathogen; fungal infection; vulvovaginitis.

XX Candida albicans.

XX WO200034481-A2.

XX 15-JUN-2000.

XX 06-DEC-1999; 99WO-EPO09833.

XX 04-DEC-1998; 98EP-00204122.

XX (JANC) JANSSEN PHARM NV.

XX Contreras RH, Nelissen B, De Backer MD, Luyten WHML, Viaene JE;

PI Logghe MG, Vialard JE;

XX WPI; 2000-431302/37.

DR N-PSDB; AAA52791.

XX Novel nucleic acid molecule and polypeptides essential for survival and
 PT growth of yeast candida albicans useful for treating candida albicans
 PT associated diseases and for identifying antifungal compounds.

XX Claim 11; Fig 16; 112pp; English.

CC The present sequence is a protein which is essential for the survival and
CC growth of *Candida albicans*. This fungus causes infection, such as
CC vulvovaginitis, in humans, particularly in those who are
CC immunocompromised. The protein and its gene can be used to diagnose
CC infection, and they can be used as targets for inhibiting the
CC proliferation of the fungus. This protein and gene are particularly
CC useful as they are thought to be species-specific
XX
SQ Sequence 699 AA;

Query Match 50.0%; Score 168; DB 3; Length 699;
Best Local Similarity 53.6%; Pred. No. 1.4e-13;
Matches 37; Conservative 8; Mismatches 22; Indels 2; Gaps 2;

Qy 1 WFGGRNPACDYHLGNISRLSNKHFOI-LXXXXXXL-LNDISTNGTWTNGQVKVNSNOLL 58
Db 62 WYFGRDPNSDLQVASSRSRISNKHFOIWLNFNDKSLWIKDTSTNGTLNLSRLVKGSNYLL 121
Qy 59 SQGDEITVG 67
Db 122 NQGEIIVG 130

RESULT 8
ABP73798

ID ABP73798 standard; protein; 699 AA.

XX
AC ABP73798;

XX
DT 30-JAN-2003 (first entry)

XX
DE *Candida albicans* essential protein SEQ ID NO 7635.

XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
XX signal transduction; DNA replication; cell division; growth;
XX proliferation; *Candida albicans*; fungicide; antifungal.

XX
OS *Candida albicans*.

XX
PN WO200253728-A2.

XX
PD 11-JUL-2002.

XX
PF 26-DEC-2001; 2001WO-US049486.

XX
PR 29-DEC-2000; 2000US-0259128P.

XX
PR 20-FEB-2001; 2001US-00792024.

XX
PR 22-AUG-2001; 2001US-0314050P.

XX
PA (ELIT-) ELITRA PHARM INC.

XX
PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

XX
DR WPI; 2002-566694/60.

XX
DR N-PSDB; ABZ32348.

XX
PT Constructing strains for identifying gene products as effective targets

XX for therapeutic intervention, by inactivating in the strain one allele of

XX a gene and placing other allele of the gene under conditional expression.

XX
PS Claim 44; SEQ ID NO 7635; 167pp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal
XX cells in which both alleles of a gene are modified, comprising modifying
XX one allele by insertion or replacement by a cassette having an
XX expressible selectable marker and modifying other allele by
XX recombination, of a promoter replacement fragment with a heterologous
XX promoter, so that expression of the second allele is regulated by the
XX promoter. (M1) is useful for constructing a strain of diploid fungal
XX cells in which both alleles of a gene are modified. The diploid fungal
XX cells having both alleles modified are useful for identifying a gene that
XX is essential to the survival or growth of a fungus, a gene that
XX contributes to the virulence and/or pathogenicity of a fungus, a gene

CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of *C. albicans* cells and for
CC treating infection by *C. albicans*. The present sequence is that of an
CC essential *Candida albicans* protein used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office
XX
SQ Sequence 699 AA;

Query Match 50.0%; Score 168; DB 5; Length 699;
Best Local Similarity 53.6%; Pred. No. 1.4e-13;
Matches 37; Conservative 8; Mismatches 22; Indels 2; Gaps 2;

Qy 1 WFGGRNPACDYHLGNISRLSNKHFOI-LXXXXXXL-LNDISTNGTWTNGQVKVNSNOLL 58
Db 62 WYFGRDPNSDLQVASSRSRISNKHFOIWLNFNDKSLWIKDTSTNGTLNLSRLVKGSNYLL 121
Qy 59 SQGDEITVG 67
Db 122 NQGEIIVG 130

RESULT 9
ADS44244

ID ADS44244 standard; protein; 609 AA.

XX
AC ADS44244;

XX
DT 02-DEC-2004 (first entry)

XX
DE Bacterial polypeptide #22674.

XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.

XX
OS Bacteria.

XX
PN US2003233675-A1.

XX
PD 18-DEC-2003.

XX
PF 20-FEB-2003; 2003US-00369493.

XX
PR 21-FEB-2002; 2002US-0360039P.

XX
PA (CAOY/) CAO Y.

XX
PA (HINK/) HINKLE G J.

XX
PA (SLAT/) SLATER S C.

XX
PA (CHEN/) CHEN X.

XX
PA (GOLD/) GOLDMAN B S.

XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX
DR WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX
PS Claim 1; SEQ ID NO 22674; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 609 AA;

Query Match 27.7%; Score 93; DB 8; Length 609;
 Best Local Similarity 23.2%; Pred. No. 0.0018;
 Matches 33; Conservative 12; Mismatches 21; Indels 76; Gaps 4;
 QY 1 WTEGRNACDYHLGNISRLSNKHQFI-----LXXXXXLLNDI----- 38
 Db 60 WRFRHKSCEVL-NGPRVSNFHEIYQVYKIIIVASXTRLVLELOSLDIIIXRNDPGH 118
 QY 39 -----STNGTWN----- 46
 Db 119 RNDSESENVVFLHGKSKGLKNLLKSSASTFSNNFRHSSNGTFLNPFVBSLRLQFTDV 178
 QY 47 -GQKVEKNQNLSSQGDIEITVG 67
 Db 179 IXERLAKNSRTILSNGDEIRIG 200

RESULT 10
 ADN19646
 ID ADN19646 standard; protein; 474 AA.
 XX
 AC ADN19646;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #2299.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX

PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.

CAO Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 2299; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 474 AA;

Query Match 25.7%; Score 86.5; DB 8; Length 474;
 Best Local Similarity 23.3%; Pred. No. 0.01;
 Matches 23; Conservative 14; Mismatches 25; Indels 7; Gaps 2;

QY 4 GRNPACDYHLGNISRLSNKHQFIILXXXXXLLN-----DISTNGTWNQKVEKNQNL 57
 Db 78 GRSNTCNVQLLOFT-ASYKHFRVYVLIIDDMDPVLYCEDQSSNGTFLNHRLKGKNSVL 136

QY 58 LSQGDIEITV 66
 Db 137 LSDGDILDV 145

RESULT 11
 ABO68456
 ID ABO68456 standard; protein; 558 AA.
 XX
 AC ABO68456;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #631.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 XX US6551795-B1.
 XX
 XX 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.
XX DR N-PSDB; ABD02027.
XX PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 17202; 455pp; English.
XX CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-*P. aeruginosa* drugs, as templates for recombinant
XX CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX CC infection, and in detection of *P. aeruginosa* sequences or other sequences
XX CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 558 AA;

Query Match 25.4%; Score 85.5; DB 7; Length 558;
Best Local Similarity 33.3%; Pred. No. 0.017;
Matches 23; Conservative 13; Mismatches 30; Indels 3; Gaps 2;

QY 2 TFGNPACDYHLGNISRL-SNKHFOILXXXXXLLNDISTGTWL--NGQVKVNSQLL 58
Db 92 TIGRGPNDWLPDPERLVSSRHCTILNRDGVVYLTDTSTNGVLLVNAHRLRGNSEPL 151
QY 59 SQGDEITVG 67
Db 152 QDGETVLG 160

RESULT 12
AAU15856
ID AAU15856 standard; protein; 426 AA.
AC AAU15856;
XX 07-NOV-2001 (first entry)
XX Human novel secreted protein, Seq ID 809.
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulneryary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
OS Homo sapiens.
XX

PT An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
PS Example 2; SEQ ID NO 501; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
XX
SQ Sequence 623 AA;

Query Match 25.3%; Score 85; DB 5; Length 623;
Best Local Similarity 34.8%; Pred. No. 0.023;
Matches 23; Conservative 6; Mismatches 35; Indels 2; Gaps 1;

QY 1 WTEGRNPACDYHLGNISRLSNKHFOIL--XXXXXXLLNDISTNGTWLNGOKVEKNSQLL 58
Db 38 WTIGRRRGCDLSPFSNKLVS GDHCRIVVDEKSGQVLTEDTSTGTVINKLKVVKKQTCPL 97

QY 59 SQGDEI 64
Db 98 QTGDVI 103

RESULT 19
AAB93168
ID AAB93168 standard; protein; 652 AA.
XX AAB93168;
AC AAB93168;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:12100.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
PN EP1074617-A2.
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR
XX 27-AUG-1999; 99JP-00300253.
PR
XX 11-JAN-2000; 2000JP-00118776.
PR
XX 02-MAY-2000; 2000JP-00183767.
PR
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
PS
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 652 AA;

Query Match 25.3%; Score 85; DB 4; Length 652;
Best Local Similarity 34.8%; Pred. No. 0.025;
Matches 23; Conservative 6; Mismatches 35; Indels 2; Gaps 1;

QY 1 WTEGRNPACDYHLGNISRLSNKHFOIL--XXXXXXLLNDISTNGTWLNGOKVEKNSQLL 58
Db 38 WTIGRRRGCDLSPFSNKLVS GDHCRIVVDEKSGQVLTEDTSTGTVINKLKVVKKQTCPL 97

QY 59 SQGDEI 64
Db 98 QTGDVI 103

RESULT 20
AAB83843
ID AAB83843 standard; protein; 664 AA.
XX AAB83843;
AC AAB83843;
XX
DT 22-AUG-2001 (first entry)
DE Amino acid sequence of a human ring finger protein designated FHAR1.
XX FHAR1; RING finger protein; cancer; vaccine.
XX
XX Homo sapiens.
XX
XX WO200142430-A1.
PN
XX
XX 14-JUN-2001.
PD
XX
XX 07-DEC-2000; 2000WO-US033094.
XX
XX 08-DEC-1999; 99US-00456876.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX
XX Zhou B, Zhu Y, Chaturvedi P, Hurle MR, Li X;
PI
XX
XX WPI; 2001-381663/40.
DR
XX N-PSDB; AAF89709.
DR
XX
XX New FHAR1 polypeptide, a member of the RING finger protein family for
PT diagnosing and treating cancer, and for use in anti-cancer vaccines.
PT
XX
XX Claim 1; Page 19; 28pp; English.
PS
XX The present sequence represents a FHAR1 polypeptide, which is a member of
CC the RING finger protein family. FHAR1 is useful in the treatment of
CC cancer, and as a vaccine for inducing an immunological response in a

CC be used to detect inhibitors of cell cycle regulatory factors which can
CC be applied in the development of drugs for treating cancers and
CC proliferative diseases. Cds1 is strongly expressed in the testis as well
CC as other tissues. Cds1 can efficiently phosphorylate cdc25 and histone H1
XX
SQ Sequence 543 AA;

Query Match 25.0%; Score 84; DB 3; Length 543;
Best Local Similarity 29.1%; Pred. No. 0.027;
Matches 23; Conservative 11; Mismatches 31; Indels 14; Gaps 3;

QY 1 WTGFRNPACYD-----HLGNISRLSNKHFQIL-----XXXXXLLNDISNGTWLNG 47
| | | | | : | : | | | | | : | | | | : | | | : |
Db 114 W-FGRDKSCYCFDEPLLKRTGYTKXKHFRIKRVGPKNSYIAYIEDHSGNGTFVNT 172
| | | | | : | : | | | | | : | | | | : | | | : |

QY 48 QKVEKNSNQLLSQDEITV 66
| | | | | : | : | | | | | : | | | | : | | | : |
Db 173 ELVGKGRRLNNNSEIAL 191
| | | | | : | : | | | | | : | | | | : | | | : |

RESULT 23
AAG89798
ID AAG89798 standard; protein; 288 AA.
XX
AC AAG89798;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 3552.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH65017.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 3552; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 288 AA;

Query Match 24.3%; Score 81.5; DB 4; Length 288;
Best Local Similarity 38.5%; Pred. No. 0.027;
Matches 25; Conservative 11; Mismatches 24; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSNKHFQILXXXXXXXXLLNDI-STNGTWLNGOKVEKNSNQLLSQGD 62
| | | | | : | : | | | | | : | | | | : | | | : |
Db 217 GRNSDADLRLPD-TGVSROHVEITWDRDAIILVDLAKSTNGTTVNDTPVD---NWLADGD 272
| | | | | : | : | | | | | : | | | | : | | | : |

QY 63 EITVG 67
| | | | |
Db 273 VITVG 277
| | | | |

RESULT 24
AAW32352
ID AAW32352 standard; protein; 148 AA.
XX
AC AAW32352;
XX
DT 13-JAN-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen TbrA10.
XX
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9709429-A2.
XX
PD 13-MAR-1997.
XX
PF 30-AUG-1996; 96WO-US014675.
XX
PR 01-SEP-1995; 95US-00523435.
PR 22-SEP-1995; 95US-00532136.
PR 22-MAR-1996; 96US-00620280.
PR 05-JUN-1996; 96US-00658800.
PR 12-JUL-1996; 96US-00680573.
XX
XX (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX
DR WPI; 1997-192904/17.
DR N-PSDB; AAT91402.
XX
PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
PT useful for diagnosis of M. tuberculosis infection.
XX
XX Example 3; Page 100-101; 190pp; English.
XX
XX A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC TbrA10. The immunogenic polypeptide can be used to diagnose
CC M.tuberculosis infection by forming complexes with specific antibodies in
CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be
CC used as diagnostic primers or probes and agents that bind to the antigen,
CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
CC also used for diagnosis
XX
SQ Sequence 148 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.022;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFQILXXXXXXXXLLNDI-STNGTWLNGOKVEKNS 54
| | | | | : | : | | | | | : | | | | : | | | : |


```

Db      66 GRHPSDIFLDDVTYSRRHAERLENNEFNVV-----DVGSLNGTYVNPVD--- 111
QY      55 NQLLSQGDEITVG 67
       : :| | |:|
Db      114 SAVLANGDEVQIG 126

RESULT 25
ID AAW32420
AC AC AAW32420;
XX XX
DT 08-JAN-1998 (first entry)
XX XX
DE Mycobacterium tuberculosis antigen TBra10.
XX XX
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX XX
OS Mycobacterium tuberculosis.
XX XX
PN WO9709428-A2.
XX XX
PD 13-MAR-1997.
XX XX
PF 30-AUG-1996; 96WO-US014674.
XX XX
PR 01-SEP-1995; 95US-00523436.
XX PR 22-SEP-1995; 95US-00533634.
XX PR 22-MAR-1996; 96US-00620874.
XX PR 05-JUN-1996; 96US-00659683.
XX PR 12-JUL-1996; 96US-00680574.
XX XX
PA (CORI-) CORIXA CORP.
XX XX
PI Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX XX
WPI: 1997-192903/17.
DR DR N-PSDB; AAT91465.
XX XX
PT New immunogenic polypeptide(s). from Mycobacterium tuberculosis - are
PT useful in vaccines for prevention or treatment of tuberculosis, also for
PT diagnosis.
XX XX
PS Example 3; Page 94; 168pp; English.
XX XX
CC A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC TBra10. The immunogenic protein, and fusion proteins containing one or
CC more of the proteins or one of the proteins plus ESAT-6, are useful in
CC vaccines, preferably when formulated with a non-specific adjuvant, to
CC induce an immune response against M.tuberculosis (for treatment or
CC prevention)
XX XX
SQ Sequence 148 AA;

Query Match          23.7%; Score 79.5; DB 2; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.022;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4

QY      4 GRNPACDYHLGNIS-----RLSNKHFIQLXXXXXXXLNDI-STGTWLNGQKVEKNS 54
       ||:|||::|
Db      66 GRHPSDIFLDDVTYSRRHAERLENNEFNVV-----DVGSLNGTYVNPVD--- 113
QY      55 NQLLSQGDEITVG 67
       : :| | |:|
Db      114 SAVLANGDEVQIG 126

```

```

AC AAW81655;
XX
XX 27-JAN-1999 (first entry)
XX
XX M. tuberculosis immunogenic polypeptide TbrA10.
XX
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
XX vaccine; pharmaceutical; infection; diagnosis.
XX
XX Mycobacterium tuberculosis.
XX
XX WO9816646-A2.
XX
XX 23-APR-1998.
XX
XX 07-OCT-1997; 97WO-US018293.
XX
XX 11-OCT-1996; 96US-00730510.
XX
XX 13-MAR-1997; 97US-00818112.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TS, Twardzik DR, Lodes MJ;
XX WPI; 1998-261042/23.
XX
XX N-PSDB; AAV64448.
XX
XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
XX develop products for the detection of M. tuberculosis infection and for
XX diagnosis, treatment and prevention of tuberculosis.
XX
XX Example 3; Page 96; 230pp; English.
XX
XX This sequence represents an immunogenic portion of a soluble
XX Mycobacterium tuberculosis (MT) antigen which can be used in a method for
XX inducing protective immunity against tuberculosis (TB). This sequence can
XX be formulated into vaccines and/or pharmaceutical compositions for
XX immunising against M. tuberculosis infection or may be used for the
XX diagnosis of tuberculosis
XX
XX SQ Sequence 148 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.022;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQVKEKNS 54
DB 66 GRHPDSDFLDDVTVSRRAHFLENNEFNVV-----DVGSLNGTYVNRPEVD--- 113
QY 55 NQLLSQGEITVG 67
DB 114 SAVLANGDEVQIG 126

RESULT 29
AAV38957
ID AAY38957 standard; protein; 148 AA.
AC
XX
XX 05-NOV-1999 (first entry)
XX
XX M. tuberculosis antigen TbrA10 amino acid sequence.
XX
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
XX immunotherapy; diagnosis; immunisation; vaccine; infection;
XX immune response; skin test.
XX
XX Mycobacterium tuberculosis.
XX
XX WO9942076-A2.
XX
XX 26-AUG-1999.
XX
XX 17-FEB-1999; 99WO-US003268.
XX
XX 18-FEB-1998; 98US-00025197.
XX
XX 05-MAY-1998; 98US-00072967.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX WPI; 1999-527409/44.
XX

```

```

PD 26-AUG-1999.
XX
XX 17-FEB-1999; 99WO-US003265.
XX
XX 18-FEB-1998; 98US-00024753.
XX
XX 05-MAY-1998; 98US-00072596.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX WPI; 1999-527416/44.
XX
XX N-PSDB; AAZ19038.
XX
XX New polypeptide comprising antigenic portions of M. tuberculosis.
XX
XX Example 3; Page 136; 323pp; English.
XX
XX This invention describes novel recombinant antigens and their encoding
XX nucleic acids derived from Mycobacterium tuberculosis. The novel
XX polypeptides are useful for detecting M. tuberculosis infection in a
XX biological sample by detecting antibodies which bind with the
XX polypeptides, and are useful as vaccines for immunizing against M.
XX tuberculosis infection. The new detection methods are needed as current
XX vaccination strategies do not provide 100% immunity
XX
XX SQ Sequence 148 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.022;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQVKEKNS 54
DB 66 GRHPDSDFLDDVTVSRRAHFLENNEFNVV-----DVGSLNGTYVNRPEVD--- 113
QY 55 NQLLSQGEITVG 67
DB 114 SAVLANGDEVQIG 126

RESULT 29
AAV39094
ID AAY39094 standard; protein; 148 AA.
XX
XX
XX AAY39094;
XX
XX 05-NOV-1999 (first entry)
XX
XX M. tuberculosis antigen TbrA10 amino acid sequence.
XX
XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
XX immunotherapy; diagnosis; immunisation; vaccine; infection;
XX immune response; skin test.
XX
XX Mycobacterium tuberculosis.
XX
XX WO9942076-A2.
XX
XX 26-AUG-1999.
XX
XX 17-FEB-1999; 99WO-US003268.
XX
XX 18-FEB-1998; 98US-00025197.
XX
XX 05-MAY-1998; 98US-00072967.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX WPI; 1999-527409/44.
XX

```

DR N-PSDB; AAZ19250.

XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin

PT tests and protective or therapeutic vaccines or compositions.

XX

PS Example 3; Page 96-97; 299pp; English.

XX

CC The present invention describes polypeptides comprising an immunogenic

CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are

CC vaccines and fusion protein containing M. tuberculosis Ag's. M.

CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other

CC polypeptides fragments, can be used in pharmaceutical compositions or

CC vaccines to generate a protective or therapeutic immune response to M.

CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.

CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural

CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249

CC to AAZ19460 and AAZ19083 to AAZ19225 are used in the exemplification of

CC the present invention

XX

SQ Sequence 148 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 148;

Best Local Similarity 28.8%; Pred. No. 0.022;

Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHQFQILXXXXXXLLNDI-STNGTWLNGQKVEKNS 54

DB 66 GRHPDSDIFLDDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTYVNRBPVD--- 113

QY 55 NQLLSQGEITVG 67

DB 114 SAVLANGDEVQIG 126

RESULT 30

AAW72886

ID AAW72886 standard; protein; 161 AA.

XX

AC AAW72886;

XX

DT 21-JAN-1999 (first entry)

XX

DE Mycobacterium tuberculosis antigen CFP17.

XX

KW Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;

KW infection.

XX

OS Mycobacterium tuberculosis.

XX

PN W09844119-A1.

XX

PD 08-OCT-1998.

XX

PF 01-APR-1998; 98WO-DK000132.

XX

PR 02-APR-1997; 97DK-00000376.

PR 18-APR-1997; 97US-0044624P.

PR 10-NOV-1997; 97DK-00001277.

PR 05-JAN-1998; 98US-0070488P.

XX

PA (STAT-) STATENS SERUM INST.

XX

PI Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB;

PI Oettinger T, Florio W;

XX

DR WPI; 1998-542705/46.

DR N-PSDB; AAW63917.

XX

XX New isolated mycobacteria polypeptides and nucleic acids - used for

PT developing products for the diagnosis of or vaccination against

PT mycobacterial infections, particularly tuberculosis.

XX

PS Claim 1; Page 127-128; 163pp; English.

XX The present sequence represents a Mycobacterium tuberculosis protein.

CC Products from the present invention, which describes protein fragments

CC and nucleic acid fragments derived from M. tuberculosis, can be used in

CC the detection of and prevention of mycobacterial infections. In

CC particular, the proteins and nucleic acids can be used for the diagnosis

CC of or vaccination against tuberculosis caused by M. tuberculosis, M.

CC africanum or M. bovis

XX

SQ Sequence 161 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 161;

Best Local Similarity 28.8%; Pred. No. 0.024;

Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHQFQILXXXXXXLLNDI-STNGTWLNGQKVEKNS 54

DB 79 GRHPDSDIFLDDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTYVNRBPVD--- 126

QY 55 NQLLSQGEITVG 67

DB 127 SAVLANGDEVQIG 139

RESULT 31

AAZ1903

ID AAZ1903 standard; protein; 162 AA.

XX

AC AAZ1903;

XX

DT 06-SEP-1999 (first entry)

XX

DE Amino acid sequence of antigen CFP17.

XX

KW Immunogenic; Mycobacterium tuberculosis; immune response; infection;

KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;

KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;

KW CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;

KW CFP25A; CFP30B; CFP7B.

XX

OS Mycobacterium tuberculosis.

XX

PN W09924577-A1.

XX

PD 20-MAY-1999.

XX

PF 08-OCT-1998; 98WO-DK000438.

XX

PR 10-NOV-1997; 97DK-00001277.

PR 05-JAN-1998; 98US-0070488P.

PR 01-APR-1998; 98WO-DK000132.

XX

PA (STAT-) STATENS SERUM INST.

XX

PI Andersen P, Skjot R;

XX

DR WPI; 1999-347282/29.

DR N-PSDB; AAX81003.

XX

PT New immunogenic fragment of Mycobacterium tuberculosis.

XX

XX Example 3; Page 59; 265pp; English.

XX The invention describes a substantially pure immunogenic polypeptide

CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a

CC protective immune response against infections by mycobacteria belonging

CC to the tuberculosis complex. The invention provides a (1) fusion

CC polypeptide comprising at least one polypeptide fragment (I) and at least

CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell

CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second

CC different amino acid sequence from M. tuberculosis, and/or including a

CC sequence which protects the first amino acid sequence from in vivo

CC degradation or post-translational processing; (3) a nucleic acid fragment

CC that encodes the above polypeptides. The polypeptides and nucleic acid
 CC are useful as pharmaceuticals, for diagnosis of and as antigens for
 CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
 CC bovis. The polypeptides are also useful for diagnosing ongoing or
 CC tuberculosis infection in an animal with bacteria belonging to the
 CC CFP30A or a T-cell epitope of for the induction of a strong immune
 CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
 CC epitope for diagnosis of TB in a mammal by performing a DTH type skin
 CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
 CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
 CC cell epitope of for the preparation of an immunological composition; and
 CC for the preparation of a subunit vaccine
 XX
 SQ Sequence 162 AA;

Query Match 23.7%; Score 79.5; DB 2; Length 162;
 Best Local Similarity 28.8%; Pred. No. 0.024;
 Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
 DB 80 GRHPDSDFLDDVTVSRRHAEFRLENNEFNVV-----DVGSLNGTYVNRPEVD--- 127

QY 55 NQLLSQSGDEITVG 67

DB 128 SAVLANGDEVQIG 140

RESULT 32

ID AAM50735 standard; protein; 162 AA.

XX AAM50735;

DT 18-APR-2002 (first entry)

DE Mycobacterium tuberculosis immunodominant Mtb protein CP17.

XX Mtb; CP17; immunogen; mycobacteria; immunisation; vaccine.

XX Mycobacterium tuberculosis.

XX WO200204018-A2.

XX 17-JAN-2002.

XX 10-JUL-2001; 2001WO-US021717.

XX 10-JUL-2000; 2000US-0217646P.

PA (COLS) UNIV COLORADO STATE RES FOUND.

PI Orme IM, Belisle JT;

XX WPI; 2002-164602/21.

XX Vaccine for boosting immunity to mycobacteria when administered in mid-
 PT life in a subject who has been vaccinated in childhood with Bacillus
 PT Calmette-Guerrin, has purified proteins from mycobacterium tuberculosis.

PS Claim 8; Page 17; 61pp; English.

XX The present sequence is that of the Mycobacterium tuberculosis (Mtb)
 CC strain H37Rv gene Rv1827 product, designated CP17. This is one of 31
 CC immunodominant secreted or cytosolic Mtb proteins of strain H37Rv (see
 CC AAM50729-59) discovered through the use of 2-dimensional liquid phase
 CC electrophoresis coupled with an in vitro interferon-gamma assay and
 CC liquid chromatography-mass spectrometry. The immunogens stimulate a
 CC strong interferon-gamma response from T cells of M. tuberculosis infected
 CC mice. The invention provides vaccine compositions for boosting immunity
 CC to mycobacteria when administered in mid-life to a subject who has been
 CC vaccinated neonatally or in early childhood with BCG and in whom

CC protective immunity has waned. The vaccine compositions comprise 1 or
 CC more of the 31 purified immunogenic proteins. When used as immunogens,
 CC the secreted Mtb proteins lack the secreted signal sequence. A preferred
 CC protein is Ag85A (see AAM50759), the secreted product of the Rv3084v gene
 XX
 SQ Sequence 162 AA;

Query Match 23.7%; Score 79.5; DB 5; Length 162;
 Best Local Similarity 28.8%; Pred. No. 0.024;
 Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
 DB 80 GRHPDSDFLDDVTVSRRHAEFRLENNEFNVV-----DVGSLNGTYVNRPEVD--- 127

QY 55 NQLLSQSGDEITVG 67

DB 128 SAVLANGDEVQIG 140

RESULT 33

ID AAM48994 standard; protein; 514 AA.

XX AAM48994;

DT 08-JUL-2002 (first entry)

XX Human Chk2 phosphoenzyme inhibitor.

XX Human; Chk2 phosphoenzyme inhibitor; cytostatic.

XX Homo sapiens.

XX JP2001346588-A.

XX 18-DEC-2001.

XX 08-JUN-2000; 2000JP-00172273.

XX 08-JUN-2000; 2000JP-00172273.

XX (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.

XX WPI; 2002-145186/19.

XX N-PSDB; AAL44749.

XX New gene encoding a protein for inhibiting human Chk2 phosphoenzyme
 PT activity.

XX Claim 2; Page 13-16; 36pp; Japanese.

XX The present invention provides the protein and coding sequences of a
 CC human Chk2 phosphoenzyme inhibitor. The sequences can be used in the
 CC chemotherapy of cancers. The present sequence is the protein of the
 CC invention

XX Sequence 514 AA;

Query Match 22.9%; Score 77; DB 5; Length 514;
 Best Local Similarity 27.8%; Pred. No. 0.23;
 Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;

QY 1 WTGRNPACDY-----HLGNISRLSNKHFOIL-----XXXXXXLLNDISTNGTWLNG 47
 DB 114 W-FGRDKSCYCFDEPLKRTDYRTYRSKKHFRIFREVGPKNSYIAYIEDHSGNGTFVNT 172

QY 48 QKVEKNSQLLSQSGDEITV 66

DB 173 ELVGKRRPLNNNSEIAL 191

RESULT 34


```
QY      1  WTFGRNPACYD-----HLGNISRLSNKHQFQL-----XXXXXXXXLLNDISTNGTWLNG 47
ID      AAY06204
DB      113  W-FGRDKSCYCFDEPLLRKTDKYRTYSKGFHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 171
QY      48  QKVEKNSQLLSQGDITV 66
DB      172  ELVGKGRPLNNSEIAL 190

RESULT 36
AAY06204
ID      AAY06204 standard; protein; 543 AA.
AC      AAY06204;
XX
XX      16-AUG-1999 (first entry)
XX      Human checkpoint kinase hCDS1.
DE      Checkpoint kinase; hCDS1; human; DNA damage; proliferative disease;
KW      cancer; therapy.
XX      Homo sapiens.
XX      WO9925843-A2.
XX      27-MAY-1999.
XX      21-OCT-1998; 98WO-EP006981.
XX      22-OCT-1997; 97GB-00022320.
XX      (SCRI ) SCRIPPS RES INST.
XX      (JANC ) JANSSEN PHARM NV.
XX      Luyten WHML, Parker AE, McGowan C, Blasina A;
XX      WPI; 1999-338010/28.
XX      N-PSDB; AAX58793.
XX      New human kinase used for treatment of cancer and proliferative diseases.
XX      Claim 9; Fig 2; 39pp; English.
XX      The present sequence represents a novel human checkpoint kinase, termed
CC      hCDS1, that acts in coordination with Cdc25 in the DNA damage checkpoint.
CC      The sequence is predicted from a cDNA (see AAX58793) isolated from a SK-N
CC      -MC neuroblastoma cDNA library. The predicted protein is 28% identical to
CC      the cdc1 protein of S. pombe. A 2.2 kb transcript is expressed in testis
CC      and in 8 human cancer samples examined. hCDS1, its inhibitors and
CC      activators, are useful for treating cancer or proliferative disease
CC      (claimed). Inhibitors and activators of the kinase activity can also be
CC      used in anti-cancer therapy, particularly by increasing susceptibility of
CC      cancer cells to chemotherapy and/or radiotherapy (claimed). hCDS1 is
CC      useful for modifying DNA damage checkpoint activity of a cancer cell
CC      (claimed)
XX
XX      Sequence 543 AA;
QY      Query Match 22.9%; Score 77; DB 2; Length 543;
      Best Local Similarity 27.8%; Pred. No. 0.24;
      Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;
QY      1  WTFGRNPACYD-----HLGNISRLSNKHQFQL-----XXXXXXXXLLNDISTNGTWLNG 47
DB      114  W-FGRDKSCYCFDEPLLRKTDKYRTYSKGFHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 172
QY      48  QKVEKNSQLLSQGDITV 66
DB      173  ELVGKGRPLNNSEIAL 191
```

```
RESULT 37
AAY05765
ID      AAY05765 standard; protein; 543 AA.
XX      AAY05765;
AC      AAY05765;
XX      02-AUG-1999 (first entry)
XX      Novel human checkpoint kinase hCDS1.
DE      Cell cycle checkpoint kinase; human; hCDS1; cell proliferation; cancer;
KW      therapy; chemotherapy; adjunct.
XX      Homo sapiens.
XX      WO9920747-A2.
XX      29-APR-1999.
XX      21-OCT-1998; 98WO-EP006982.
XX      22-OCT-1997; 97GB-00022320.
XX      (JANC ) JANSSEN PHARM NV.
XX      Luyten WHML, Parker AE;
XX      WPI; 1999-288300/24.
XX      N-PSDB; AAX25464.
XX      Novel human cell cycle checkpoint kinase hCDS1, useful for treating cell
PT      proliferation diseases, e.g. cancer.
XX      Claim 2; Page 33-34; 35pp; English.
XX      The present sequence represents hCDS1, a novel human cell cycle
CC      checkpoint kinase that acts in coordination with Cdc25 at the DNA damage
CC      checkpoint, rather than the replication checkpoint as found in yeast. The
CC      hCDS1 sequence was predicted from hCDS1 cDNA (see AAX25464) that was
CC      obtained from a human SK-N-MC neuroblastoma cDNA library. It shows 28%
CC      identity with cdc1 protein of S. pombe, 28% identity to RAD53 and 27%
CC      identity to the DUN1 Kinase of S. cerevisiae. Northern blot analysis
CC      identified a single transcript of about 2.2 kb expressed in testis and in
CC      8 human cancer samples examined. The characterisation of hCDS1 and the
CC      elucidation of its role in the DNA damage checkpoint allows for the
CC      preparation of pharmaceuticals and therapeutic methods for acting as an
CC      adjunct to chemotherapy of cancer. Thus, pharmaceutical formulations
CC      incorporating hCDS1 cDNA, RNA, antisense molecules, hCDS1 protein,
CC      antibodies against the protein, or other therapeutics identified in
CC      assays of the invention, can be administered in conjunction with any
CC      suitable chemotherapy agent to act as an adjunct to the main action of
CC      the chemotherapy agent
XX
XX      Sequence 543 AA;
QY      Query Match 22.9%; Score 77; DB 2; Length 543;
      Best Local Similarity 27.8%; Pred. No. 0.24;
      Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;
QY      1  WTFGRNPACYD-----HLGNISRLSNKHQFQL-----XXXXXXXXLLNDISTNGTWLNG 47
DB      114  W-FGRDKSCYCFDEPLLRKTDKYRTYSKGFHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 172
QY      48  QKVEKNSQLLSQGDITV 66
DB      173  ELVGKGRPLNNSEIAL 191

RESULT 38
AAM48996
ID      AAM48996 standard; protein; 543 AA.
XX      AAM48996;
AC      AAM48996;
```

XX 08-JUL-2002 (first entry)
PT Human Chk2 phosphoenzyme inhibitor related protein #2.
XX Human; Chk2 phosphoenzyme inhibitor; cytostatic.
XX Homo sapiens.
XX JP2001346588-A.
XX 18-DEC-2001.
XX 08-JUN-2000; 2000JP-00172273.
XX 08-JUN-2000; 2000JP-00172273.
XX (IGAK-) IGAU SEIBUTSUGAKU KENKYUSHO KK.
XX WPI; 2002-145186/19.
XX N-PSDB; AAL44751.
XX New gene encoding a protein for inhibiting human Chk2 phosphoenzyme activity.
XX Disclosure; Page 27-30; 36pp; Japanese.
XX The present invention provides the protein and coding sequences of a human Chk2 phosphoenzyme inhibitor. The sequences can be used in the chemotherapy of cancers. The present sequence is a protein described in the exemplification of the invention
XX Sequence 543 AA;
XX Query Match 22.9%; Score 77; DB 5; Length 543;
XX Best Local Similarity 27.8%; Pred. No. 0.24;
XX Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;
XX 1 WTGFRNPACDY-----HLGNISRLSNKHQFQIL-----XXXXXXLLNDISTNGTWLNG 47
XX 114 W-FGRDKSCYCFDEPLLRKTRDKYRTYSKKHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 172
XX 48 QKVEKNSQLLSQGEITV 66
XX 173 ELVGKGRRLNNSEIAL 191
XX RESULT 39
XX AAM48995
XX ID AAM48995 standard; protein; 543 AA.
XX AC AAM48995;
XX DT 08-JUL-2002 (first entry)
XX DE Human Chk2 phosphoenzyme inhibitor related protein #1.
XX KW Human; Chk2 phosphoenzyme inhibitor; cytostatic.
XX OS Homo sapiens.
XX PN JP2001346588-A.
XX PD 18-DEC-2001.
XX PF 08-JUN-2000; 2000JP-00172273.
XX PR 08-JUN-2000; 2000JP-00172273.
XX PA (IGAK-) IGAU SEIBUTSUGAKU KENKYUSHO KK.
XX WPI; 2002-145186/19.
XX N-PSDB; AAL44750.
XX

XX New gene encoding a protein for inhibiting human Chk2 phosphoenzyme activity.
XX Disclosure; Page 20-23; 36pp; Japanese.
XX The present invention provides the protein and coding sequences of a human Chk2 phosphoenzyme inhibitor. The sequences can be used in the chemotherapy of cancers. The present sequence is a protein described in the exemplification of the invention
XX Sequence 543 AA;
XX Query Match 22.9%; Score 77; DB 5; Length 543;
XX Best Local Similarity 27.8%; Pred. No. 0.24;
XX Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;
XX 1 WTGFRNPACDY-----HLGNISRLSNKHQFQIL-----XXXXXXLLNDISTNGTWLNG 47
XX 114 W-FGRDKSCYCFDEPLLRKTRDKYRTYSKKHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 172
XX 48 QKVEKNSQLLSQGEITV 66
XX 173 ELVGKGRRLNNSEIAL 191
XX RESULT 40
XX ABG30701
XX ID ABG30701 standard; protein; 543 AA.
XX AC ABG30701;
XX DT 07-OCT-2002 (first entry)
XX DE Human checkpoint kinase 2 (CHK2) polypeptide.
XX KW Human; checkpoint kinase 2; CHK2; enzyme; antisense therapy.
XX OS Homo sapiens.
XX PN WO200251858-A2.
XX PD 04-JUL-2002.
XX PF 17-DEC-2001; 2001WO-US048966.
XX PR 22-DEC-2000; 2000US-00746043.
XX PA (ISIS-) ISIS PHARM INC.
XX PA (ABBO) ABBOTT LAB.
XX PI Sarthy A, Cowsett LM;
XX WPI; 2002-575367/61.
XX N-PSDB; ABK88909.
XX New antisense oligonucleotides targeted to a nucleic acid encoding checkpoint kinase 2 (CHK2), useful for treating a disease or condition associated with CHK2, or in distinguishing functions of members of a biological pathway.
XX Disclosure; Page 87-90; 100pp; English.
XX The invention relates to an antisense compound targeted to a nucleic acid molecule encoding human checkpoint kinase 2 (CHK2). The antisense compound specifically hybridizes with and inhibits the expression of human CHK2. The antisense compounds are useful as research reagents and diagnostics, in distinguishing between functions of various members of a biological pathway, and in the treatment of a disease or disorder, which can be treated by modulating the expression of CHK2. This sequence represents the human CHK2 protein
XX Sequence 543 AA;
XX

Query Match 22.9%; Score 77; DB 5; Length 543;
Best Local Similarity 27.8%; Pred. No. 0.24;
Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;
QY 1 WFGGRNPACDY-----HLGNISRLSNKHFQIL-----XXXXXXXXLLNDISTNGTWING 47
Db 114 W-FGRDKSCYCFDEPLLKRTDKYRTYSKKHFRIFREVGPKNSYIAVIEDHSGNGTFVNT 172
QY 48 QKVEKNSNQLLSQGDEITV 66
Db 173 ELVGKGRRPENNNSEIAL 191

Search completed: February 24, 2005, 15:19:39
Job time : 143.537 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:14:52 ; Search time 33.7724 Seconds
(without alignments)
148.094 Million cell updates/sec

Title: US-09-936-956-19

Perfect score: 336

Sequence: 1 WTFGRNPACDYHLGNISRLS.....QKVEKNSQLLSQDEITVG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	315	93.8	65	4	US-09-300-008B-46
2	313.5	93.3	776	1	US-08-198-446B-17
3	313.5	93.3	776	2	US-08-870-693-17
4	313.5	93.3	821	1	US-08-198-446B-6
5	313.5	93.3	821	2	US-08-870-693-6
6	165	49.1	513	4	US-09-248-796A-18515
7	140.5	41.8	64	4	US-09-300-008B-44
8	94.5	28.1	103	4	US-09-902-540-13360
9	94.5	28.1	595	4	US-09-902-540-16334
10	93.5	27.8	311	4	US-09-902-540-10544
11	87.5	26.0	264	4	US-09-902-540-10597
12	87	25.9	476	4	US-09-902-540-14917
13	86	25.6	614	4	US-09-902-540-14372
14	85.5	25.4	558	4	US-09-252-991A-17202
15	84	25.0	323	4	US-09-902-540-10166
16	83	24.7	149	4	US-09-248-796A-18516
17	83	24.7	356	4	US-09-902-540-16059
18	79.5	23.7	148	3	US-08-818-112-64
19	79.5	23.7	148	3	US-08-818-111-65
20	79.5	23.7	148	3	US-09-056-556-64
21	79.5	23.7	148	3	US-09-072-596-65
22	79.5	23.7	148	4	US-09-072-967-64
23	79.5	23.7	162	4	US-09-050-739-6
24	78	23.2	63	4	US-09-300-008B-43
25	77	22.9	490	4	US-09-949-016-10788
26	77	22.9	543	4	US-09-529-093A-2
27	77	22.9	543	4	US-09-529-154-2

28 76.5 22.8 459 4 US-09-902-540-10583 Sequence 10583, A
29 75.5 22.5 640 4 US-09-902-540-15375 Sequence 15375, A
30 73.5 21.9 517 4 US-09-849-617-2 Appli Sequence 2, Appli
31 72 21.4 200 4 US-09-902-540-15221 Sequence 15221, A
32 69.5 20.7 236 4 US-09-902-540-16276 Sequence 16276, A
33 68.5 20.4 64 4 US-09-300-008B-45 Sequence 45, Appl
34 67 19.9 3256 4 US-09-919-172-98 Sequence 98, Appl
35 67 19.9 3256 4 US-09-976-594-22 Sequence 22, Appl
36 67 19.9 3256 4 US-09-919-039-21 Sequence 21, Appl
37 65.5 19.5 893 4 US-09-248-796A-20980 Sequence 20980, A
38 65 19.3 64 4 US-09-300-008B-41 Sequence 41, Appl
39 64 19.0 63 4 US-09-300-008B-42 Sequence 42, Appl
40 64 19.0 177 4 US-09-902-540-12792 Sequence 12792, A
41 63 18.8 440 4 US-09-198-452A-755 Sequence 755, App
42 63 18.8 845 4 US-09-438-185A-714 Sequence 714, App
43 62.5 18.6 476 4 US-09-252-991A-22693 Sequence 22693, A
44 62 18.5 65 4 US-09-300-008B-48 Sequence 48, Appl
45 62 18.5 176 4 US-09-902-540-11336 Sequence 11336, A

ALIGNMENTS

RESULT 1

US-09-300-008B-46

; Sequence 46, Application US/093000008B

; Patent No. 6458534

; GENERAL INFORMATION:

; APPLICANT: Concannon et al.

; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE

; FILE REFERENCE: 9924-0003-228

; CURRENT APPLICATION NUMBER: US/09/300,008B

; PRIOR FILING DATE: 1999-04-27

; PRIOR APPLICATION NUMBER: US 60/083,269

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 46

; LENGTH: 65

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-300-008B-46

Query Match 93.8%; Score 315; DB 4; Length 65;
Best Local Similarity 91.0%; Pred. No. 6.6e-36;
Matches 61; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 1 WTFGRNPACDYHLGNISRLSFKHFIQILXXXXXLLNDISTNGTWLNGKVEKNSQLLSQ 60

Db 1 WTFGRNPACDYHLGNISRLSFKHFIQIL- -LGNLLNDISTNGTWLNGKVEKNSQLLSQ 58

Qy 61 GDEITVG 67

Db 59 GDEITVG 65

RESULT 2

US-08-198-446B-17

; Sequence 17, Application US/08198446B

; Patent No. 5674996

; GENERAL INFORMATION:

; APPLICANT: Hartwell, Leland H.

; APPLICANT: Weinert, Ted A.

; APPLICANT: Plon, Sharon E.

; APPLICANT: Grondine, Mark T.

; TITLE OF INVENTION: Cell Cycle Checkpoint Genes

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Christensen O'Connor Johnson Kindness PLLC

; STREET: 1420 Fifth Ave., Suite 2800

; CITY: Seattle

; STATE: WA

```
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHC17537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Yeast MEC2 protein
; US-08-198-446B-17

Query Match 93.3%; Score 313.5; DB 1; Length 776;
Best Local Similarity 89.7%; Pred. No. 2.2e-34;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTEGRNPACDYHLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQQLLS 59
Db 66 WTEGRNPACDYHLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQQLLS 125

Qy 60 QGDEITVG 67
Db 126 QGDEITVG 133

RESULT 3
US-08-870-693-17
; Sequence 17, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,693
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,446
; FILING DATE: February 18, 1994
; APPLICATION NUMBER: PCT/US93/04458
; FILING DATE: May 12, 1993
; APPLICATION NUMBER: US 07/884,426
```

```
; FILING DATE: May 14, 1992
; APPLICATION NUMBER: US 07/882,051
; FILING DATE: May 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHC110798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Yeast MEC2 protein
; US-08-870-693-17

Query Match 93.3%; Score 313.5; DB 2; Length 776;
Best Local Similarity 89.7%; Pred. No. 2.2e-34;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WTEGRNPACDYHLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQQLLS 59
Db 66 WTEGRNPACDYHLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQQLLS 125

Qy 60 QGDEITVG 67
Db 126 QGDEITVG 133

RESULT 4
US-08-198-446B-6
; Sequence 6, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHC17537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 821 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

DESCRIPTION: Yeast MEC2 protein
US-08-198-446B-6

Query Match 93.3%; Score 313.5; DB 1; Length 821;
Best Local Similarity 89.7%; Pred. No. 2.3e-34;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WFGFNPCADYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNAGQKVEKSNOLL 59
|||||
Db 66 WFGFNPCADYHLGNISRLSNKHFOILLGDEGNLLNDISTNGTWNAGQKVEKSNOLL 125
|||||

Qy 60 QGDEITVG 67
|||||
Db 126 QGDEITVG 133
|||||

RESULT 5
US-08-870-693-6
; Sequence 6, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/870,693
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,446
; FILING DATE: February 18, 1994
; APPLICATION NUMBER: PCT/US93/04458
; FILING DATE: May 12, 1993
; APPLICATION NUMBER: US 07/884,426
; FILING DATE: May 14, 1992
; APPLICATION NUMBER: US 07/882,051
; FILING DATE: May 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI10798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 821 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Yeast MEC2 protein
US-08-870-693-6

Query Match 93.3%; Score 313.5; DB 2; Length 821;
Best Local Similarity 89.7%; Pred. No. 2.3e-34;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WFGFNPCADYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNAGQKVEKSNOLL 59
|||||
Db 66 WFGFNPCADYHLGNISRLSNKHFOILLGDEGNLLNDISTNGTWNAGQKVEKSNOLL 125
|||||

Qy 60 QGDEITVG 67
|||||
Db 126 QGDEITVG 133
|||||

RESULT 6
US-09-248-796A-18515
; Sequence 18515, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18515
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18515

Query Match 49.1%; Score 165; DB 4; Length 513;
Best Local Similarity 52.2%; Pred. No. 3.2e-14;
Matches 36; Conservative 9; Mismatches 22; Indels 2; Gaps 2;

Qy 1 WFGFNPCADYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWNAGQKVEKSNOLL 58
|||||
Db 91 WYFGDPNSDLQVASSKISNKHFOIWNFNKSLWKIDSTNGTHLNSRLVKGSNYLL 150
|||||

Qy 59 SQGDEITVG 67
|||||
Db 151 NQGDIEAVG 159
|||||

RESULT 7
US-09-300-008B-44
; Sequence 44, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-09-300-008B-44

Query Match 41.8%; Score 140.5; DB 4; Length 64;
Best Local Similarity 46.3%; Pred. No. 6e-12;
Matches 31; Conservative 13; Mismatches 20; Indels 3; Gaps 2;

Qy 1 WFGFNPCADYHLGNISRLSNKHFOILLXXXXXXLLNDISTNGTWNAGQKVEKSNOLL 60
|||||
Db 1 WGFGRHKSCEVNL-NGPRVSNFPEI--YQGVFLHDHSSNGTFLNFKAKNSRTILSN 57
|||||

US-09-902-540-14917
; Sequence 14917, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14917
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14917

Query Match 25.9%; Score 87; DB 4; Length 476;
Best Local Similarity 36.9%; Pred. No. 0.0015;
Matches 24; Conservative 10; Mismatches 27; Indels 4; Gaps 2;
Qy 4 GRNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNQLLSQGD 62
Db 69 GAHPGNDLVLDVETASRRHFEIQTTERGYLLVDLSTNGTFLDGRRIER---AYLSPGS 125
Qy 63 EITVG 67
Db 126 QIRAG 130

RESULT 13
US-09-902-540-14372
; Sequence 14372, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14372
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14372

Query Match 25.6%; Score 86; DB 4; Length 614;
Best Local Similarity 34.8%; Pred. No. 0.0029;
Matches 24; Conservative 16; Mismatches 21; Indels 8; Gaps 4;
Qy 2 TFGNPACDYHL--GNISRLSNKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNQLLS 58
Db 43 TIGRLPGNDIILAKGNVSKY---HSRIVAKDGRFIIVDMKSTNGTFVNGRKIA--APQVL 97
Qy 59 SQGDEITVG 67
Db 98 KPTDQVYIG 106

RESULT 14
US-09-902-540-14372
; Sequence 14372, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14372
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14372

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17202
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17202

Query Match 25.4%; Score 85.5; DB 4; Length 558;
Best Local Similarity 33.3%; Pred. No. 0.003;
Matches 23; Conservative 13; Mismatches 30; Indels 3; Gaps 2;
Qy 2 TFGNPACDYHLGNISRL-SNKHFIQLXXXXXXLLNDISTNGTWL--NGQKVEKNSNQLLS 58
Db 92 TIGRGPNDMDVLPDPERLVSSRHCTILNRDGVVYLLTDTSTNGVLLVNAVHRLRGNSPGL 151
Qy 59 SQGDEITVG 67
Db 152 QDGETVRLG 160

RESULT 15
US-09-902-540-10166
; Sequence 10166, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10166
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10166

Query Match 25.0%; Score 84; DB 4; Length 323;
Best Local Similarity 39.7%; Pred. No. 0.0025;
Matches 29; Conservative 9; Mismatches 19; Indels 16; Gaps 5;
Qy 2 TFGNPACDYH-----LGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
Db 76 TIGR----DQHNHIVVDLNVSR---RHARILGRGGKMLVEDLGTSTNGTFLNDQEVQLAS 128
Qy 55 NQLLSQGDEITVG 67
Db 129 P--LRSGDLVKVG 139

RESULT 16
US-09-248-796A-18516
; Sequence 18516, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN


```
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-818-111-65

Query Match 23.7%; Score 79.5; DB 3; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.004;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
Db 66 GRHPDSIFLDDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTYVNRFPVD--- 113

QY 55 NQLLSQGEITVG 67
Db 114 SAVLANGDEVQIG 126

RESULT 20
US-09-056-556-64
; Sequence 64, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-072-596-65

Query Match 23.7%; Score 79.5; DB 4; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.004;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
Db 66 GRHPDSIFLDDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTYVNRFPVD--- 113

; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-818-111-65

Query Match 23.7%; Score 79.5; DB 3; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.004;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

QY 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
Db 66 GRHPDSIFLDDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTYVNRFPVD--- 113

QY 55 NQLLSQGEITVG 67
Db 114 SAVLANGDEVQIG 126

RESULT 20
US-09-056-556-64
; Sequence 64, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
```

QY	55	NQLLSQGDDEITVG 67
Db	114	SAVLANGDEVQIG 126
RESULT 22		
US-09-072-967-64		
; Sequence 64, Application US/09072967		
; Patent No. 6592877		
; GENERAL INFORMATION:		
; APPLICANT: Reed, Steven G.		
; APPLICANT: Skeiky, Yasir A.W.		
; APPLICANT: Dillon, Davin C.		
; APPLICANT: Campos-Neto, Antonio		
; APPLICANT: Houghton, Raymond		
; APPLICANT: Vedwick, Thomas S.		
; APPLICANT: Twardzik, Daniel R.		
; APPLICANT: Lodes, Michael J.		
; APPLICANT: Hendricksen, Ronald C.		
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY		
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS		
; NUMBER OF SEQUENCES: 355		
; CORRESPONDENCE ADDRESS:		
; ADDRESSEE: SEED AND BERRY LLP		
; STREET: 6300 Columbia Center, 701 Fifth Avenue		
; CITY: Seattle		
; STATE: Washington		
; COUNTRY: USA		
; ZIP: 98104-7092		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: PatentIn Release #1.0, Version #1.30		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/09/072,967		
; FILING DATE: 05-MAY-1998		
; CLASSIFICATION:		
; ATTORNEY/AGENT INFORMATION:		
; NAME: Maki, David J.		
; REGISTRATION NUMBER: 31,392		
; REFERENCE/DOCKET NUMBER: 210121.411C9		
; TELECOMMUNICATION INFORMATION:		
; TELEPHONE: (206) 622-4900		
; TELEFAX: (206) 682-6031		
; INFORMATION FOR SEQ ID NO: 64:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 148 amino acids		
; TYPE: amino acid		
; STRANDEDNESS: single		
; TOPOLOGY: linear		
; US-09-072-967-64		
Query Match 23.7%; Score 79.5; DB 4; Length 148;		
Best Local Similarity 28.8%; Pred. No. 0.004;		
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;		
QY	4	GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQKVKNS 54
Db	66	GRHPDSDFLDDVTVSRRAEFLNENFNVV-----DVGSLNGTYVNRPEVD--- 113
QY	55	NQLLSQGDDEITVG 67
Db	114	SAVLANGDEVQIG 126
RESULT 23		
US-09-050-739-6		
; Sequence 6, Application US/09050739		
; Patent No. 6641814		
; GENERAL INFORMATION:		
; APPLICANT: ANDERSEN, Peter		
; APPLICANT: RASMUSSEN, Peter Birk		
; APPLICANT: ROSENKRANDS, Ida		
; APPLICANT: WELDRING, Karin		
; APPLICANT: FLORIO, Walter		
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS		
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS		
; FILE REFERENCE: 670001-2002.1		
; CURRENT APPLICATION NUMBER: US/09/050,739		
; CURRENT FILING DATE: 1998-03-30		
; EARLIER FILING DATE: 1997-04-02		
; EARLIER FILING DATE: 1997-11-10		
; EARLIER FILING DATE: 1997-04-18		
; EARLIER FILING DATE: 1997-04-18		
; EARLIER FILING DATE: 1998-01-05		
; NUMBER OF SEQ ID NOS: 173		
; SOFTWARE: PatentIn Ver. 2.0		
; SEQ ID NO 6		
; LENGTH: 162		
; TYPE: PRP		
; ORGANISM: Mycobacterium tuberculosis		
; US-09-050-739-6		
Query Match 23.7%; Score 79.5; DB 4; Length 162;		
Best Local Similarity 28.8%; Pred. No. 0.0044;		
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;		
QY	4	GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXLLNDI-STNGTWLNGQKVKNS 54
Db	80	GRHPDSDFLDDVTVSRRAEFLNENFNVV-----DVGSLNGTYVNRPEVD--- 127
QY	55	NQLLSQGDDEITVG 67
Db	128	SAVLANGDEVQIG 140
RESULT 24		
US-09-300-008B-43		
; Sequence 43, Application US/09300008B		
; Patent No. 6458534		
; GENERAL INFORMATION:		
; APPLICANT: Concannon et al.		
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE		
; TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE		
; FILE REFERENCE: 9924-0003-228		
; CURRENT APPLICATION NUMBER: US/09/300,008B		
; CURRENT FILING DATE: 1999-04-27		
; PRIOR APPLICATION NUMBER: US 60/083,269		
; PRIOR FILING DATE: 1998-04-27		
; NUMBER OF SEQ ID NOS: 64		
; SOFTWARE: FastSeq for Windows Version 3.0		
; SEQ ID NO 43		
; LENGTH: 63		
; TYPE: PRP		
; ORGANISM: Saccharomyces cerevisiae		
; US-09-300-008B-43		
Query Match 23.2%; Score 78; DB 4; Length 63;		
Best Local Similarity 35.7%; Pred. No. 0.0023;		
Matches 25; Conservative 12; Mismatches 21; Indels 12; Gaps 5		
QY	2	TFGRNPACDYHLG--NISRLSNKHFOILXXXXXXLLN--DISTNGTWLNGQKVKNS 57
Db	2	TIGRSRCDVILSEPDIPTF-HAEFHL-----LINVIDKSRNGTFINGNELVK-KDYI 53
QY	58	LSQGDDEITVG 67
Db	54	LKNGDRIVFG 63


```
RESULT 25
US-09-949-016-10788
; Sequence 10788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10788
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10788

Query Match          22.9%; Score 77; DB 4; Length 490;
Best Local Similarity 27.8%; Pred. No. 0.037;
Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;

Qy 1 WTCGRNPACDY-----HLGNISRLSNKHQFQL-----XXXXXXLLNDISTNGTWLNG 47
Db 114 W-FGRDKSCYCFDEPLLRKTDKYRTYKXKHFRIFREVGPKNSYIAVIEDHSGNGTFTVNT 172
Qy 48 QKVEKSNQLLSQGEITV 66
Db 173 ELVGKGRPLNNSEIAL 191

RESULT 26
US-09-529-093A-2
; Sequence 2, Application US/09529093A
; Patent No. 6413755
; GENERAL INFORMATION:
; APPLICANT: LUYTEN, Walter H.M.L.
; APPLICANT: PARKER, Andrew E.
; APPLICANT: MCGOWAN, Clare H.
; APPLICANT: BLASINA, Alessandra
; TITLE OF INVENTION: Human Checkpoint Kinase, hCDS1, Compositions and Methods
; FILE REFERENCE: TSRI 649.0
; CURRENT APPLICATION NUMBER: US/09/529,093A
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/EP98/06981
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: GB 9722320.0
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-093A-2

Query Match          22.9%; Score 77; DB 4; Length 543;
Best Local Similarity 27.8%; Pred. No. 0.042;
Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;

Qy 1 WTCGRNPACDY-----HLGNISRLSNKHQFQL-----XXXXXXLLNDISTNGTWLNG 47
Db 114 W-FGRDKSCYCFDEPLLRKTDKYRTYKXKHFRIFREVGPKNSYIAVIEDHSGNGTFTVNT 172
Qy 48 QKVEKSNQLLSQGEITV 66
Db 173 ELVGKGRPLNNSEIAL 191

RESULT 27
US-09-529-154-2
; Sequence 2, Application US/09529154
; Patent No. 6531312
; GENERAL INFORMATION:
; APPLICANT: LUYTEN, Walter H.M.L.
; APPLICANT: PARKER, Andrew E.
; TITLE OF INVENTION: A NOVEL HUMAN CHECKPOINT KINASE, hCDS1, COMPOSITIONS
; FILE REFERENCE: JAB 1333 (US)
; CURRENT APPLICATION NUMBER: US/09/529,154
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/EP98/06982
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: GB 9722320.0
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-154-2

Query Match          22.9%; Score 77; DB 4; Length 543;
Best Local Similarity 27.8%; Pred. No. 0.042;
Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;

Qy 1 WTCGRNPACDY-----HLGNISRLSNKHQFQL-----XXXXXXLLNDISTNGTWLNG 47
Db 114 W-FGRDKSCYCFDEPLLRKTDKYRTYKXKHFRIFREVGPKNSYIAVIEDHSGNGTFTVNT 172
Qy 48 QKVEKSNQLLSQGEITV 66
Db 173 ELVGKGRPLNNSEIAL 191

RESULT 28
US-09-902-540-10583
; Sequence 10583, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10583
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10583

Query Match          22.8%; Score 76.5; DB 4; Length 459;
Best Local Similarity 30.8%; Pred. No. 0.04;
Matches 20; Conservative 17; Mismatches 23; Indels 5; Gaps 3;

Qy 4 GRNPACDYHLGNISRLSNKHQFQLXXXXXXXXLLNDI-STNGTWLNGQKVEKSNQLLSQGD 62
Db 50 GSDSACDLVLT-DATVSRHLEVERTARGMLRDTGSRNGTFLDGRQVQVQ---AYLTSGD 105
Qy 63 BITVG 67
Db 63 BITVG 67
```

```
Db      106 KVELG 110

RESULT 29
US-09-902-540-15375
; Sequence 15375, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; CURRENT FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15375
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15375

Query Match      22.5%; Score 75.5; DB 4; Length 640;
Best Local Similarity 31.8%; Pred. No. 0.083;
Matches 21; Conservative 10; Mismatches 18; Indels 17; Gaps 2;

Qy      16 ISRLSNKH-----FOILXXXXXXLLNDISTNGTWLNGOKVEKNSNQLLSQG 61
Db      28 IGRLLSAHLCLEDEKVSRIHSVIEASDGAMSIIDMGVSGTYVNGKRVTKG---LLSFG 84
Qy      62 DEITVG 67
Db      85 DEIRVG 90

RESULT 30
US-09-849-617-2
; Sequence 2, Application US/09849617
; Patent No. 6593110
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CIT1350-1
; CURRENT APPLICATION NUMBER: US/09/849,617
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-849-617-2

Query Match      21.9%; Score 73.5; DB 4; Length 517;
Best Local Similarity 25.3%; Pred. No. 0.12;
Matches 20; Conservative 12; Mismatches 34; Indels 13; Gaps 2;

Qy      1 WFGFNPACDY-----HLGNISRLSNKHQFQL-----XXXXXXLLNDISTNGTWLNG 47
Db      85 YVFGDRKKCDYDFIPVLNQTRDYKTSKRHRIFQELGHGHSRVANIEDLSGNGTFVNK 144
Qy      48 QKVEKNSNQLLSQGEITV 66
Db      145 EIIKGRTLPUNNAEIAL 163

RESULT 31
US-09-902-540-15221
; Sequence 15221, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; CURRENT FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15221
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15221

Query Match      21.4%; Score 72; DB 4; Length 200;
Best Local Similarity 33.3%; Pred. No. 0.061;
Matches 22; Conservative 12; Mismatches 26; Indels 6; Gaps 3;

Qy      4 GRNPACDYHLGNISRLSNKHQFI--LXXXXXXLLNDISTNGTWLNGOKVEKNSNQLLSQG 61
Db      50 GRSESAQFQVED-DGISRKHAKVVALDGRFQLVDLASTNGTYLNGLKVNAP---LYDG 105
Qy      62 DEITVG 67
Db      106 DKIQIG 111

RESULT 32
US-09-902-540-16276
; Sequence 16276, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; CURRENT FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16276
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16276

Query Match      20.7%; Score 69.5; DB 4; Length 236;
Best Local Similarity 38.0%; Pred. No. 0.17;
Matches 19; Conservative 9; Mismatches 19; Indels 3; Gaps 2;

Qy      19 LSNKHQFIILXXXXXXLLNDI-STNGTWLNGOKVEKNSNQLLSQGEITV 67
Db      109 VSRTHARVLHGEFGSLEDLGSFNGVFLNGQRVD--SRVPLADGDVIEMG 156

RESULT 33
US-09-300-008B-45
; Sequence 45, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
```

; TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9224-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-300-008B-45

Query Match 20.4%; Score 68.5; DB 4; Length 64;
Best Local Similarity 34.4%; Pred. No. 0.047;
Matches 21; Conservative 9; Mismatches 28; Indels 3; Gaps 2;
Qy 4 GRNPACDYHLGNISRLSNKHQFQILXXXXXXLLNDISTNGTWNLGQKVEKNSNQLLSQGE 63
Db 4 GRNDKQVLTPNS-ISSVH--CVFWCMVFYKDCSLNGTYLNGLLKRDKTYLLKHCDV 60
Qy 64 I 64
Db 61 I 61

RESULT 34
US-09-919-172-98
; Sequence 98, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 98
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 2700132CD1
US-09-919-172-98

Query Match 19.9%; Score 67; DB 4; Length 3256;
Best Local Similarity 30.8%; Pred. No. 8.8;
Matches 20; Conservative 13; Mismatches 28; Indels 4; Gaps 3;
Qy 3 FGPNPACDYHLGNISRLSNKHQFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQGE 61
Db 29 FGRGIECDIRI-QLPVVVSKQKIEIHEQEAHLNFSTNPTQVNGSVIDEFVR--LKHG 85
Qy 62 DEITV 66
Db 86 DVITI 90

RESULT 35
US-09-976-594-22
; Sequence 22, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US

; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 2700132CD1
US-09-976-594-22

Query Match 19.9%; Score 67; DB 4; Length 3256;
Best Local Similarity 30.8%; Pred. No. 8.8;
Matches 20; Conservative 13; Mismatches 28; Indels 4; Gaps 3;
Qy 3 FGPNPACDYHLGNISRLSNKHQFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQGE 61
Db 29 FGRGIECDIRI-QLPVVVSKQKIEIHEQEAHLNFSTNPTQVNGSVIDEFVR--LKHG 85
Qy 62 DEITV 66
Db 86 DVITI 90

RESULT 36
US-09-919-039-21
; Sequence 21, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 3256
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 2700132CD1
US-09-919-039-21

Query Match 19.9%; Score 67; DB 4; Length 3256;
Best Local Similarity 30.8%; Pred. No. 8.8;
Matches 20; Conservative 13; Mismatches 28; Indels 4; Gaps 3;
Qy 3 FGPNPACDYHLGNISRLSNKHQFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQGE 61
Db 29 FGRGIECDIRI-QLPVVVSKQKIEIHEQEAHLNFSTNPTQVNGSVIDEFVR--LKHG 85
Qy 62 DEITV 66
Db 86 DVITI 90

RESULT 37
US-09-248-796A-20980
; Sequence 20980, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132

```
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20980
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20980

Query Match          19.5%; Score 65.5; DB 4; Length 883;
Best Local Similarity 27.5%; Pred. No. 2.9;
Matches 19; Conservative 16; Mismatches 23; Indels 11; Gaps 2;

Qy      8  ACDYHLGNISRLSNKHFOIL-----XXXXXXLLNDISTNGTTLNGQKVEKNSNQLLS 59
Db      82  ATNYHPLDVTKLSNEKVSIGISKQNRVTVSSEKLFDDSGDITELNGVNLKNTSKQTLN 141
Qy      60  Q---GDEIT 65
Db      142  EFDGDDLT 150

RESULT 38
US-09-300-008B-41
; Sequence 41, Application US/093000008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-300-008B-41

Query Match          19.3%; Score 65; DB 4; Length 64;
Best Local Similarity 28.8%; Pred. No. 0.14;
Matches 17; Conservative 10; Mismatches 30; Indels 2; Gaps 1;

Qy      9  CDYHLGNISRLSNKHFOILXXXXXXLLNDISTNGTTLNGQKVEKNSNQLLSQGDDEITVG 67
Db      8  CAILIENDQSISRNH--AVLTANVLTLDKNSKYGTFFVNEEKMQNGFSGRTLKSGDGIITFG 64

RESULT 39
US-09-300-008B-42
; Sequence 42, Application US/093000008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 63

; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20980
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-300-008B-42

Query Match          19.0%; Score 64; DB 4; Length 63;
Best Local Similarity 32.8%; Pred. No. 0.19;
Matches 21; Conservative 10; Mismatches 29; Indels 4; Gaps 3;

Qy      3  FGRNPACDYHLGNISRLSNKHFOILXXXXXXLLNDISTNGTTLNGQKVEKNSNQLLSQGD 62
Db      3  FGRGIECDIRI-QLPVVSQKHCKI-FIHEAILHNFSTNFTQVNGSVIDEFVR--LKHGD 58
Qy      63  EITV 66
Db      59  VITI 62

RESULT 40
US-09-902-540-12792
; Sequence 12792, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12792
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12792

Query Match          19.0%; Score 64; DB 4; Length 177;
Best Local Similarity 34.5%; Pred. No. 0.66;
Matches 10; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy      39  STNGTTLNGQKVEKNSNQLLSQGDDEITVG 67
Db      125  SMNGTWNASSLIGPDQDRMLNDGDALAFG 153

Search completed: February 24, 2005, 15:25:31
Job time : 34.7724 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:23:51 ; Search time 96.9594 Seconds
(without alignment)
226.127 Million cell updates/sec

Title: US-09-936-956-19

Perfect score: 336

Sequence: 1 WTEGRNPACDYHLGNISRLS.....QKVEKNSQLLSQDEITVG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	313.5	93.3	68	17	US-10-857-622-56
2	313.5	93.3	522	9	US-09-740-627-11
3	313.5	93.3	821	14	US-10-081-119-16
4	168	50.0	699	14	US-10-032-585-7635
5	137.5	40.9	460	9	US-09-740-627-10
6	93	27.7	609	15	US-10-369-493-22674
7	89	26.5	190	15	US-10-424-599-154547
8	86.5	25.7	174	15	US-10-369-493-22999
9	85	25.3	426	9	US-09-764-864-809
10	85	25.3	664	9	US-09-780-525-2
11	84	25.0	543	9	US-09-740-627-1
12	81.5	24.3	288	9	US-09-738-626-3552
13	80	23.8	1141	16	US-10-437-963-156334

14	79.5	23.7	148	14	US-10-193-002-65	Sequence 65, Appl
15	79.5	23.7	148	14	US-10-084-843-64	Sequence 64, Appl
16	79.5	23.7	162	9	US-09-791-171-6	Sequence 6, Appl
17	79.5	23.7	162	16	US-10-620-246-6	Sequence 7, Appl
18	79.5	23.7	162	16	US-10-332-512A-7	Sequence 14509, A
19	78	23.2	324	14	US-10-156-761-14509	Sequence 2, Appl
20	77	22.9	543	14	US-10-185-182A-2	Sequence 56100, A
21	77	22.9	1082	15	US-10-425-114-56100	Sequence 50922, Ap
22	76.5	22.8	143	9	US-09-738-626-5092	Sequence 187236,
23	76.5	22.8	229	15	US-10-424-599-187236	Sequence 9, Appl
24	76.5	22.8	459	9	US-09-740-627-9	Sequence 6, Appl
25	75	22.3	545	14	US-10-142-356-6	Sequence 1578, Ap
26	74.5	22.0	162	10	US-09-804-980-6	Sequence 2, Appl
27	74	22.0	513	15	US-10-369-493-1578	Sequence 2, Appl
28	73.5	21.9	517	9	US-09-849-617-2	Sequence 2, Appl
29	73.5	21.9	517	15	US-10-618-173-2	Sequence 3551, Ap
30	73	21.7	154	9	US-09-738-626-3551	Sequence 65407, A
31	70	20.8	617	15	US-10-425-114-65407	Sequence 56757, A
32	69	20.5	228	15	US-10-425-114-56757	Sequence 163818,
33	69	20.5	1484	16	US-10-437-963-163818	Sequence 233512,
34	68.5	20.4	376	15	US-10-424-599-233512	Sequence 13999, A
35	68.5	20.4	843	14	US-10-156-761-13999	Sequence 9696, Ap
36	68	20.2	212	14	US-10-156-761-9696	Sequence 4, Appl
37	67.5	20.1	488	14	US-10-080-170-4	Sequence 4, Appl
38	67.5	20.1	488	16	US-10-080-170-4	Sequence 4, Appl
39	67.5	20.1	488	16	US-10-468-356-4	Sequence 1109, Ap
40	67	19.9	92	9	US-09-764-864-1109	Sequence 98, Appl
41	67	19.9	3256	9	US-09-919-172-98	Sequence 21, Appl
42	67	19.9	3256	10	US-09-919-039-21	Sequence 174, App
43	67	19.9	3256	16	US-10-408-765A-174	Sequence 9, Appl
44	67	19.9	3256	16	US-10-701-490-9	Sequence 430, App
45	65.5	19.5	397	15	US-10-389-647-430	

ALIGNMENTS

RESULT 1

US-10-857-622-56
; Sequence 56, Application US/10857622
; Publication No. US20050026234A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: VIOLIN, Jonathan
; APPLICANT: NEWTON, Alexandra
; APPLICANT: TSJEN, Roger
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: Emission Ratiometric Indicators of Phosphorylation By C-Kinase
; FILE REFERENCE: 39754-0891 CPC1CP2
; CURRENT APPLICATION NUMBER: US/10/857,622
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/865,291
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 594,575
; PRIOR FILING DATE: 1996-01-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: phosphoaminoacid binding domain
US-10-857-622-56

Query Match 93.3%; Score 313.5; DB 17; Length 68;
Best Local Similarity 89.7%; Pred. No. 7.7e-34;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy 1 WtGRNPACDYLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQLLS 59
Db 1 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 60
Qy 60 QGDEITVG 67
Db 61 QGDEITVG 68

RESULT 2
US-09-740-627-11
; Sequence 11, Application US/09740627
; Patent No. US20020012964A1
; GENERAL INFORMATION:
; APPLICANT: Nakanishi, Makoto
; TITLE OF INVENTION: CELL CYCLE REGULATING FACTOR
; FILE REFERENCE: 06501-071001
; CURRENT APPLICATION NUMBER: US/09/740,627
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/JP99/03350
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: JP 10/192467
; PRIOR FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-740-627-11

Query Match 93.3%; Score 313.5; DB 9; Length 522;
Best Local Similarity 89.7%; Pred. No. 7.8e-33;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WtGRNPACDYLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQLLS 59
Db 66 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 125
Qy 60 QGDEITVG 67
Db 126 QGDEITVG 133

RESULT 3
US-10-081-119-16
; Sequence 16, Application US/10081119
; Publication No. US20030045491A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Jeffers, Anne B.
; APPLICANT: Chan, Vivien W.
; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic
; FILE REFERENCE: 16932.002
; CURRENT APPLICATION NUMBER: US/10/081,119
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/289,813
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-081-119-16

Query Match 93.3%; Score 313.5; DB 14; Length 821;
Best Local Similarity 89.7%; Pred. No. 1.3e-32;
Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 WtGRNPACDYLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQLLS 59
Db 1 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 60
Qy 60 QGDEITVG 67
Db 61 QGDEITVG 68
```

```
Db 66 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 125
Qy 60 QGDEITVG 67
Db 126 QGDEITVG 133

RESULT 4
US-10-032-585-7635
; Sequence 7635, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7635
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7635

Query Match 50.0%; Score 168; DB 14; Length 699;
Best Local Similarity 53.6%; Pred. No. 2.2e-13;
Matches 37; Conservative 8; Mismatches 22; Indels 2; Gaps 2;

Qy 1 WtGRNPACDYLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQLLS 58
Db 62 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 121
Qy 59 SQGDEITVG 67
Db 122 NQGDEITVG 130

RESULT 5
US-09-740-627-10
; Sequence 10, Application US/09740627
; Patent No. US20020012964A1
; GENERAL INFORMATION:
; APPLICANT: Nakanishi, Makoto
; TITLE OF INVENTION: CELL CYCLE REGULATING FACTOR
; FILE REFERENCE: 06501-071001
; CURRENT APPLICATION NUMBER: US/09/740,627
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/JP99/03350
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: JP 10/192467
; PRIOR FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-09-740-627-10

Query Match 40.9%; Score 137.5; DB 9; Length 460;
Best Local Similarity 41.3%; Pred. No. 1.5e-09;
Matches 31; Conservative 13; Mismatches 22; Indels 9; Gaps 2;

Qy 1 WtGRNPACDYLGNISRLSNKHFIQLLGDGNLLNDISTNGTWLNGQVKVKNQLLS 52
Db 60 WtGRNHKSEVUL-NGRVSNFHFYQGRHNSDSENVVFLDHSNGTFLNFERLAK 118
Qy 53 NSNQLSQGDEITVG 67
Db 53 NSNQLSQGDEITVG 67
```

```
Db 119 NSRTILSGDEIRIG 133

RESULT 6
US-10-369-493-22674
; Sequence 22674, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22674
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(609)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-22674

Query Match 27.7%; Score 93; DB 15; Length 609;
Best Local Similarity 23.2%; Pred. No. 0.0017;
Matches 33; Conservative 12; Mismatches 21; Indels 76; Gaps 4;

Qy 1 WFGGRNACDYHLGNISRLSKHFKQI-----LXXXXXXLLNDI----- 38
Db 60 WRFGHKSCVVL-NGPRVSNFHEIYQVYKIIIVASXTRVLVLEQLSLXDIIXNDPGH 118
Qy 39 -----STNGTWIN----- 46
Db 119 RNDSESENVVFLHGKSKGLNKLKSSASTFSNNFRHSSNGTFLNFVSESRLRLQFTDV 178
Qy 47 -GQVKENSNQLSQDEITVG 67
Db 179 IXERLAKNSRTILSGDEIRIG 200

RESULT 7
US-10-424-599-154547
; Sequence 154547, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154547
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110578C.1.pep
US-10-424-599-154547

Query Match 26.5%; Score 89; DB 15; Length 190;
Best Local Similarity 32.8%; Pred. No. 0.0015;

Matches 22; Conservative 15; Mismatches 26; Indels 4; Gaps 2;
Qy 4 GRNPACDYHLGNISRLSKHFKQIIL-----XXXXXXLLNDISTNGTWINGKVKENSNQLSQ 60
Db 53 GRHPDCDMLTHPS-ISRFLQIRSNFSSRTFSLDLSSVHGTVWSGRRIEPMVSVEMKE 111
Qy 61 GDEITVG 67
Db 112 GETLRVG 118

RESULT 8
US-10-369-493-2299
; Sequence 2299, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2299
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(474)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2299

Query Match 25.7%; Score 86.5; DB 15; Length 474;
Best Local Similarity 33.3%; Pred. No. 0.0094;
Matches 23; Conservative 14; Mismatches 25; Indels 7; Gaps 2;

Qy 4 GRNPACDYHLGNISRLSKHFKQIILXXXXXXLLN-----DISTNGTWINGKVKENSNQL 57
Db 78 GRSNTCYQLQFT-ASYKHFRVYSLVLDMDPLVYCEDQSSNGTFLNHLRIGKNSVL 136
Qy 58 LSQGEITV 66
Db 137 LSDGDILDV 145

RESULT 9
US-09-764-864-809
; Sequence 809, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 809
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```


APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazov, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 156334
LENGTH: 1141
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1141)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_56010C.1.pep
US-10-437-963-156334

Query Match 23.8%; Score 80; DB 16; Length 1141;
Best Local Similarity 26.3%; Pred. No. 0.18;
Matches 20; Conservative 16; Mismatches 20; Indels 20; Gaps 3;

Qy 3 FGRNPACD-----YHLGNISRLSNKHFQILXXXXXXXXLLNDIST-NGTWLNGQKVE 51
Db 57 FGRHPECHLVDPHSVSRFHFLEVRSRQRITV-----TDLSSVHTWISGRIP 107
Qy 52 KNSNQLSQGDEITVG 67
Db 108 PNPVELTAGDVLRLG 123

RESULT 14

US-10-193-002-65
Sequence 65, Application US/10193002
Publication No. US20030135026A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TUBERCULOSIS

NUMBER OF SEQUENCES: 350

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/193,002
FILING DATE: 10-Jul-2002
CLASSIFICATION: <Unknown>
APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 148 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-193-002-65

Query Match 23.7%; Score 79.5; DB 14; Length 148;
Best Local Similarity 28.8%; Pred. No. 0.021;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;

Qy 4 GRNPACDYHLGNIS-----RLSNKHFQILXXXXXXXXLLNDI-STNGTWLNGQKVEKNS 54
Db 66 GRHPDSDFLDDVTVSRRAHFRLENNEFNVV-----DVGSLNGTYVNRFPVD--- 113

Qy 55 NQLLSQGDITVG 67

Db 114 SAVLANGDEVQIG 126

RESULT 15

US-10-084-843-64
Sequence 64, Application US/10084843
Publication No. US20030143243A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, Davin C.
Campos-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 64:

[illegible]

```

; PRIOR APPLICATION NUMBER: PCT/EP98/06981
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: GB 9722320.0
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-185-182A-2

Query Match          22.9%   Score 77; DB 14; Length 543;
Best Local Similarity 27.8%; Pred. No. 0.2;
Matches 22; Conservative 11; Mismatches 32; Indels 14; Gaps 3;

QY      1  WTFGRNPACDY-----HLGNISRLSNKHQFQIL-----KXKXKXLLNDISTNGTWLNG 47
      Db      114  W-FGRDKSCYCYDEPLLKRTDKYRTYSKKHFRIFREVGPKNSYAIYIEDHSGNGTFVNT 172
      QY      48  QKVEKNSQLLSQSGDEITV 66
      Db      173  ELVGKGKRRPLNNNSEIAL 191

RESULT 21
US-10-425-114-56100
; Sequence 56100, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56100
; LENGTH: 1082
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMLE01810018D04_FLI.pap
US-10-425-114-56100

Query Match          22.9%   Score 77; DB 15; Length 1082;
Best Local Similarity 31.0%; Pred. No. 0.44;
Matches 22; Conservative 16; Mismatches 23; Indels 10; Gaps 4;

QY      4  GRNPACDYHL-----GNISRLSNKHQFQILKXKXKXLLNDISTNGTWL-NGQKVEKNSQ 56
      Db      2  GSNRSCNFPNDQITISGNLCKI--KHPQ-GDGSAAVAVLESMGSKGSLVNGTHVKKNTSC 58

QY      57  LLSQSGDEITVG 67
      Db      59  VLNSGDEWVFG 69

RESULT 22
US-09-738-626-5092
; Sequence 5092, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

```

```
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5092
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5092

Query Match      22.8%; Score 76.5; DB 9; Length 143;
Best Local Similarity 31.3%; Pred. No. 0.051;
Matches 21; Conservative 15; Mismatches 26; Indels 5; Gaps 3;

QY  2  TFGNPACDYLGNISRLSNKHFOILXXXXXXLLNDI-STNGTWLNGQVKVKNQLLSQ 60
Db   69  TAGHPESDIFLDVTVTSRRHAEFRINEGEFEVVDGSLNGTYVNR---EPRNAQVMQT 124
QY  61  GDEITVG 67
Db   125  GDEIQIG 131

RESULT 23
US-10-424-599-187236
; Sequence 187236, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187236
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(229)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140086C.1.pep
US-10-424-599-187236

Query Match      22.8%; Score 76.5; DB 15; Length 229;
Best Local Similarity 37.7%; Pred. No. 0.087;
Matches 26; Conservative 9; Mismatches 31; Indels 3; Gaps 2;

QY  1  WFGNPACDYLGN--ISRLSNKHFOILXXXXXXLLNDISTNGTWLNGQVKVKNQLL 58
Db   104  YMFGRDLDCFVLEHPIISRF-HAVVQFKRSGDAYLDLGSHTGTFLNKNQVEKNTYVDL 162
QY  59  SQGEITVG 67
Db   111  KIGTNRMLKNDVDVISL 128

Query Match      22.8%; Score 76.5; DB 9; Length 459;
Best Local Similarity 30.8%; Pred. No. 0.19;
Matches 24; Conservative 14; Mismatches 27; Indels 13; Gaps 3;

QY  1  WFGNPACDYLH-----GNISRLSNKHFOILXXXXXXLLN-----DISTNGTWLNGQ 48
Db   52  FTAGRGEANDLILTLNDLPEKILTRISKVHF-IKRANCELTPVVTQDLSRNGTFVYNE 110
QY  49  KVEKNSNQLLSQGEITV 66
Db   111  KIGTNRMLKNDVDVISL 128

RESULT 25
US-10-142-356-6
; Sequence 6, Application US/10142356
; Publication No. US20030036183A1
; GENERAL INFORMATION:
; APPLICANT: Boylan, John F.
; APPLICANT: Bowers, Alex J.
; TITLE OF INVENTION: Serine-Threonine Kinase Member, h2520-40
; FILE REFERENCE: 01017/37177A
; CURRENT APPLICATION NUMBER: US/10/142,356
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,276
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-142-356-6

Query Match      22.3%; Score 75; DB 14; Length 545;
Best Local Similarity 29.1%; Pred. No. 0.37;
Matches 23; Conservative 9; Mismatches 33; Indels 14; Gaps 3;

QY  1  WFGNPACDY-----HLGNISRLSNKHFOILXXX-----XXLLNDISTNGTWLNG 47
Db   117  W-FGRDKSCBYCFDGPLLKRTDKYRTYKHKHFRIFRMGPKNVCYIVVYLEDHSGNGTFVNT 175
QY  48  QKVEKNSNQLLSQGEITV 66
Db   176  ELIGKGRCPFLSNNSIAL 194

RESULT 26
```

```
Db   163  HVGDVIRFG 171

RESULT 24
US-09-740-627-9
; Sequence 9, Application US/09740627
; Patent No. US20020012964A1
; GENERAL INFORMATION:
; APPLICANT: Nakanishi, Makoto
; TITLE OF INVENTION: CELL CYCLE REGULATING FACTOR
; FILE REFERENCE: 06501-071001
; CURRENT APPLICATION NUMBER: US/09/740,627
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/JP99/03350
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: JP 10/192467
; PRIOR FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-740-627-9

Query Match      22.8%; Score 76.5; DB 9; Length 459;
Best Local Similarity 30.8%; Pred. No. 0.19;
Matches 24; Conservative 14; Mismatches 27; Indels 13; Gaps 3;

QY  1  WFGNPACDYLH-----GNISRLSNKHFOILXXXXXXLLN-----DISTNGTWLNGQ 48
Db   52  FTAGRGEANDLILTLNDLPEKILTRISKVHF-IKRANCELTPVVTQDLSRNGTFVYNE 110
QY  49  KVEKNSNQLLSQGEITV 66
Db   111  KIGTNRMLKNDVDVISL 128

RESULT 25
US-10-142-356-6
; Sequence 6, Application US/10142356
; Publication No. US20030036183A1
; GENERAL INFORMATION:
; APPLICANT: Boylan, John F.
; APPLICANT: Bowers, Alex J.
; TITLE OF INVENTION: Serine-Threonine Kinase Member, h2520-40
; FILE REFERENCE: 01017/37177A
; CURRENT APPLICATION NUMBER: US/10/142,356
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,276
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-142-356-6

Query Match      22.3%; Score 75; DB 14; Length 545;
Best Local Similarity 29.1%; Pred. No. 0.37;
Matches 23; Conservative 9; Mismatches 33; Indels 14; Gaps 3;

QY  1  WFGNPACDY-----HLGNISRLSNKHFOILXXX-----XXLLNDISTNGTWLNG 47
Db   117  W-FGRDKSCBYCFDGPLLKRTDKYRTYKHKHFRIFRMGPKNVCYIVVYLEDHSGNGTFVNT 175
QY  48  QKVEKNSNQLLSQGEITV 66
Db   176  ELIGKGRCPFLSNNSIAL 194

RESULT 26
```

```
US-09-804-980-6
; Sequence 6, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-6

Query Match      22.2%; Score 74.5; DB 10; Length 162;
Best Local Similarity 28.8%; Pred. No. 0.11;
Matches 21; Conservative 15; Mismatches 16; Indels 21; Gaps 4;

Qy 4 GRNPACDYHLGNIS-----RLSNKHFOILXXXXXXXXLLNDI-STNGTWTWNGQKVEKNS 54
Db 80 GRHPDSIPLDVVSRRAEFLNNEFNVV-----DVGSLNGTTVNRPEVD--- 127
Qy 55 NQLLSQGDDEITVG 67
Db 128 SAVLANGDEVOIG 140

RESULT 27
US-10-369-493-1578
; Sequence 1578, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1578
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1578

Query Match      22.0%; Score 74; DB 15; Length 513;
Best Local Similarity 29.1%; Pred. No. 0.47;
Matches 23; Conservative 14; Mismatches 22; Indels 20; Gaps 4;

Qy 2 TFGNRPACDYHLG-----NISRLSNKHFOILXXXXXXXXLLN--DISTNGTWTWNGQ 48
Db 57 TIGRSRCDVLTSEPDISTFHAEPHLQMDVDFQ-----RNLIINVIDKSRNGTTFINGN 110
Qy 49 KVEKNSNQLLSQGDDEITVG 67
Db 111 RLVK-KDYILKNGDRIVFG 128

RESULT 28
US-09-849-617-2
; Sequence 2, Application US/09849617
; Publication No. US2002008392A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CITI350-1
; CURRENT APPLICATION NUMBER: US/09/849,617
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-849-617-2

Query Match      21.9%; Score 73.5; DB 9; Length 517;
Best Local Similarity 25.3%; Pred. No. 0.55;
Matches 20; Conservative 12; Mismatches 34; Indels 13; Gaps 2;

Qy 1 WTEGRNPACDY-----HLGNISRLSNKHFOIL-----XXXXXXXXLLNDISTNGTWTWNG 47
Db 85 YVFGDRKKCDYTFDIPVLNQTDYKTYSKRFRIFQELGHSRVAIEDLSGNGTFVNK 144
Qy 48 QKVEKNSNQLLSQGDDEITV 66
Db 145 EIIGKGRTLPLTNNAETAL 163

RESULT 29
US-10-618-173-2
; Sequence 2, Application US/10618173
; Publication No. US20040018603A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CITI350-1
; CURRENT APPLICATION NUMBER: US/10/618,173
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US/09/849,617
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-618-173-2

Query Match      21.9%; Score 73.5; DB 15; Length 517;
Best Local Similarity 25.3%; Pred. No. 0.55;
Matches 20; Conservative 12; Mismatches 34; Indels 13; Gaps 2;

Qy 1 WTEGRNPACDY-----HLGNISRLSNKHFOIL-----XXXXXXXXLLNDISTNGTWTWNG 47
Db 85 YVFGDRKKCDYTFDIPVLNQTDYKTYSKRFRIFQELGHSRVAIEDLSGNGTFVNK 144
Qy 48 QKVEKNSNQLLSQGDDEITV 66
Db 145 EIIGKGRTLPLTNNAETAL 163

RESULT 30
US-09-738-626-3551
; Sequence 3551, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CITI350-1
; CURRENT APPLICATION NUMBER: US/09/849,617
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-849-617-2

Query Match      21.9%; Score 73.5; DB 9; Length 517;
Best Local Similarity 25.3%; Pred. No. 0.55;
Matches 20; Conservative 12; Mismatches 34; Indels 13; Gaps 2;

Qy 1 WTEGRNPACDY-----HLGNISRLSNKHFOIL-----XXXXXXXXLLNDISTNGTWTWNG 47
Db 85 YVFGDRKKCDYTFDIPVLNQTDYKTYSKRFRIFQELGHSRVAIEDLSGNGTFVNK 144
Qy 48 QKVEKNSNQLLSQGDDEITV 66
Db 145 EIIGKGRTLPLTNNAETAL 163

RESULT 30
US-09-738-626-3551
; Sequence 3551, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CITI350-1
; CURRENT APPLICATION NUMBER: US/09/849,617
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-849-617-2

Query Match      21.9%; Score 73.5; DB 9; Length 517;
Best Local Similarity 25.3%; Pred. No. 0.55;
Matches 20; Conservative 12; Mismatches 34; Indels 13; Gaps 2;

Qy 1 WTEGRNPACDY-----HLGNISRLSNKHFOIL-----XXXXXXXXLLNDISTNGTWTWNG 47
Db 85 YVFGDRKKCDYTFDIPVLNQTDYKTYSKRFRIFQELGHSRVAIEDLSGNGTFVNK 144
Qy 48 QKVEKNSNQLLSQGDDEITV 66
Db 145 EIIGKGRTLPLTNNAETAL 163
```

```

; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3551
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3551

Query Match      21.7%; Score 73; DB 9; Length 154;
Best Local Similarity 25.4%; Pred. No. 0.16;
Matches 17; Conservative 19; Mismatches 27; Indels 4; Gaps 3;

Qy      2  TFGNPACDYHLGNISRLSNKHFOILXXXXXXLLNDI-STNGTWLNGQKVEKNSQLLSQ 60
Db      83  TMGSPECTFVGVD-DYASGNHARVFKGSEWFVEDLDSRNGTFVGGTRIDQ--PEQIAV 139

Qy      61  GDEITVG 67
Db      140  GTDIRIG 146

RESULT 31
US-10-425-114-65407
; Sequence 65407, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65407
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4763-012-B1_FLI.pep
US-10-425-114-65407

Query Match      20.8%; Score 70; DB 15; Length 617;
Best Local Similarity 36.4%; Pred. No. 2;
Matches 28; Conservative 7; Mismatches 30; Indels 12; Gaps 3;

Qy      2  TFGNPACDYHLGNISRLSNKHFOIL---XXXXXXLLNDISTNGTWLNGQKVEKNS---- 54
Db      237  TLGRVPPSDLVKLD-SEVSGKHAQINNWNGKTLKWELVDMGSLNGTFLNSQAVHHPHSAGSR 295

```

```

Qy      55  ----NQLSQGDEITVG 67
Db      296  HWGEPAPLAHGDIITLG 312

RESULT 32
US-10-425-114-56757
; Sequence 56757, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56757
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17038D04_FLI.pep
US-10-425-114-56757

Query Match      20.5%; Score 69; DB 15; Length 228;
Best Local Similarity 36.4%; Pred. No. 0.86;
Matches 28; Conservative 8; Mismatches 29; Indels 12; Gaps 4;

Qy      2  TFGNPACDYHLGNISRLSNKHFOIL---XXXXXXLLNDISTNGTWLNGQKVEKNSQL 57
Db      144  TLGRVPPSDLVKLD-SEVSGKHAQINNWNGKTLKWELVDMGSLNGTFLNSQAVHHPNAGSR 202

Qy      58  ----LSQGDEITVG 67
Db      203  HWGEPAPLAHGDIITLG 219

RESULT 33
US-10-437-963-163818
; Sequence 163818, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163818
; LENGTH: 1484
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62778C.1.pep
US-10-437-963-163818

Query Match      20.5%; Score 69; DB 16; Length 1484;
Best Local Similarity 24.1%; Pred. No. 7.2;
Matches 19; Conservative 12; Mismatches 18; Indels 30; Gaps 2;

```

Query Match 20.4%; Score 68.5; DB 14; Length 843;
Best Local Similarity 29.4%; Pred. No. 4.4; Gaps 3;
Matches 20; Conservative 16; Mismatches 27; Indels 5; Gaps 3;

QY 1 WTEGRNPACDYHLGNISRLSNKHFQILXXXXXXXXLLNDI-STNGTWNLGQKVERKNSNOLLS 59
DB 21 YTLGRPQGDITLDD-ARVSWRHATISWNGRSWVIEDHGSTNGTTFVQGORIHQWE---IG 76
QY 60 QGDEITVG 67
DB 77 PGSVAHLG 84

RESULT 36
US-10-156-761-9696
; Sequence 9696, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9696
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9696

Query Match 20.2%; Score 68; DB 14; Length 212;
Best Local Similarity 30.8%; Pred. No. 1.1; Gaps 3;
Matches 20; Conservative 14; Mismatches 27; Indels 4; Gaps 3;

QY 4 GRNPACDYHLGNISRLSNKHFQILXXXXXXXXLLNDI-STNGTWNLGQKVERKNSNOLLSQGD 62
DB 142 GRDPASGLRLSH-DTVSRVHAELSRQGLWLRDLGSTNGTTVNGRHV--IGAADVVRDGD 198
QY 63 EITVG 67
DB 199 QVSFG 203

RESULT 37
US-10-080-170-4
; Sequence 4, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495,0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Mycobacterium leprae

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:12:40 ; Search time 26.1463 Seconds
(without alignments)
246.555 Million cell updates/sec

Title: US-09-936-956-19

Perfect score: 336

Sequence: 1 WFGNRPACDVHLGNISRLS.....QKVEKNQLLSQGEITVG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1.*
- 2: Pir2.*
- 3: Pir3.*
- 4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	313.5	93.3	821	1 A39616	protein kinase RAD
2	137.5	40.9	460	2 S58862	protein kinase Cds
3	87	25.9	159	2 A96962	FHA-domain contain
4	86.5	25.7	445	2 T43420	probable protein k
5	85.5	25.4	497	2 P83634	hypothetical prote
6	81	24.1	952	2 T18837	hypothetical prote
7	80	23.8	399	2 AC3089	conserved hypotet
8	80	23.8	399	2 G98197	hypothetical prote
9	79.5	23.7	162	2 D70721	hypothetical prote
10	78.5	23.4	1524	2 A96950	DNA segregation AT
11	78	23.2	1587	2 G86457	hypothetical prote
12	75.5	22.5	234	2 AF2214	hypothetical prote
13	74.5	22.2	162	2 T44758	hypothetical prote
14	74	22.0	513	1 S43941	protein kinase DUN
15	71.5	21.3	386	2 B71407	hypothetical prote
16	70	20.8	583	2 T02045	kinase associated
17	69.5	20.7	157	2 S76488	hypothetical prote
18	69.5	20.7	445	2 AB2022	hypothetical prote
19	69.5	20.7	546	2 AB1899	adenylate cyclase
20	67.5	20.1	463	2 T10015	hypothetical prote
21	67.5	20.1	488	2 P86911	conserved hypotet
22	67	19.9	2897	2 B48666	cell proliferation
23	67	19.9	2938	2 T30249	cell proliferation
24	67	19.9	3256	2 A48666	cell proliferation
25	66.5	19.8	230	2 A83872	hypothetical prote
26	65.5	19.5	397	2 D83477	hypothetical prote
27	65	19.3	579	2 T37248	probable matrix me
28	65	19.3	598	2 T32166	hypothetical prote
29	64.5	19.2	527	2 B70700	hypothetical prote

RESULT 1

A39616

protein kinase RAD53 (EC 2.7.1.1) - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein p2588; protein YPL153c; SPK1 protein

C;Species: Saccharomyces cerevisiae

C;Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004

C;Accession: A39616; S65164; S69446; S13321

R;Stern, D.F.; Zheng, P.; Beidler, D.R.; Zerillo, C.

Mol. Cell. Biol. 11, 987-1001, 1991

A;Title: Spk1, a new kinase from Saccharomyces cerevisiae, phosphorylates proteins on se

A;Reference number: A39616; MUID:91117267; PMID:1899289

A;Accession: A39616

A;Molecule type: DNA

A;Residues: 1-821 <STE>

A;Cross-references: UNIPROT:P22216; GB:M55623; NID:gl72656; PIDN:AAA35070.1; PID:gl72657

A;Experimental source: strain S288C

R;Zheng, P.; Fay, D.S.; Burton, J.; Xiao, H.; Pinkham, J.L.; Stern, D.F.

Mol. Cell. Biol. 13, 5829-5842, 1993

A;Title: SPK1 is an essential S-phase-specific gene of Saccharomyces cerevisiae that enc

A;Reference number: A54697; MUID:93361015; PMID:8355715

A;Contents: annotation

R;Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S65154

A;Accession: S65164

A;Molecule type: DNA

A;Residues: 1-821 <PUR>

A;Cross-references: EMBL:Z73509; NID:gl370325; PIDN:CAA97858.1; PID:gl370326; GSPDB:GN001

R;Purnelle, B.; Combiez, S.; Coster, F.; Naveau, F.; Goffeau, A.

submitted to the EMBL Data Library, March 1996

A;Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies
ogue to the human phosphoryl phosphatase activator PTPA and a homologue to the plant

A;Reference number: S69428

A;Accession: S69446

A;Molecule type: DNA

A;Residues: 1-821 <PUN>

A;Cross-references: EMBL:X96770; NID:gl403537; PIDN:CAA65568.1; PID:gl403556

C;Genetics:

A;Gene: SGD:RAD53; SPK1; MEC2; SAD1; MIPS:YPL153c

A;Cross-references: SGD:S0006074; MIPS:YPL153c

A;Map position: 16L

C;Function:

A;Description: serine/threonine-specific protein kinase

A;Note: contains low activity as tyrosine-specific protein kinase

C;Superfamily: protein kinase SPK1; kinase interaction domain homology; protein kinase h

C;Keywords: ATP; cell cycle control; nucleus; phosphotransferase; serine/threonine-speci

F;68-133/Domain: kinase interaction domain homology <KIH>

F;196-466/Domain: protein kinase homology <KIN>

F;204-213/Region: protein kinase ATP-binding motif

F;227/Active site: Lys #status predicted


```
Qy 4 GRNPACDYHLGNISRLSNKHFOILXXXXXLLN-----DISTNGTWLNGQVKVNSQL 57
Db 65 GRSTNCYQLQFT-ASYKHFRVSVLIDDMDFLVCEQDSNGTFLNHRLLGKNSVL 123
Qy 58 LSQSGDEITV 66
Db 124 LSDGDILDV 132

RESULT 5
F83634
hypothetical protein PA0081 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83634
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-497 <STO>
A:Cross-references: UNIPROT:Q91751; GB:AE004447; GB:AE004091; NID:g9945902; PIDN:AAG0347
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0081

Query Match 25.4%; Score 85.5; DB 2; Length 497;
Best Local Similarity 33.3%; Pred. No. 0.0054;
Matches 23; Conservative 13; Mismatches 30; Indels 3; Gaps 2;

Qy 2 TFGRNPCADYHLGNISRL-SNKHFOILXXXXXLLNDISTNGTWL--NGQKVKVNSQL 58
Db 29 TIGRGPNDVLPDPRLVSRHCTILNRDGVVYLTDTSTNGVLLVNGHRLRGNSEPL 88
Qy 59 SQSGDEITVG 67
Db 89 QDGETVRLG 97

RESULT 6
T18837
hypothetical protein C01G6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18837
R:Berks, M.
submitted to the EMBL Data Library, August 1994
A:Reference number: Z19029
A:Accession: T18837
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-952 <WIL>
A:Cross-references: UNIPROT:P46012; EMBL:Z35595; PIDN:CAA84636.1; GSPDB:GN000020; CESP:C0
A:Experimental source: clone C01G6
C:Genetics:
A:Gene: CESP:C01G6.5
A:Map position: 2
A:Introns: 34/2; 61/2; 101/3; 195/1; 218/3; 845/3; 886/3

Query Match 24.1%; Score 81; DB 2; Length 952;
Best Local Similarity 41.2%; Pred. No. 0.042;
Matches 14; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 34 LLNDISTNGTWLNGQVKVNSQLSQSGDEITVG 67
Db 84 LVEDISENGTVINDRRLSKDKREILKSGDTIKFG 117
```

RESULT 7

```
AC3089
conserved hypothetical protein Atu4335 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AC3089
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3089
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <KUR>
A:Cross-references: UNIPROT:Q8U7W4; GB:AE008689; PIDN:AAL45129.1; PID:g17742801; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4335
A:Map position: linear chromosome

Query Match 23.8%; Score 80; DB 2; Length 399;
Best Local Similarity 26.7%; Pred. No. 0.022;
Matches 20; Conservative 16; Mismatches 31; Indels 8; Gaps 2;

Qy 1 WTF-----GRNPACDYHL-GNISRLSNKHFOILXXXXXLLNDISTNGTWLNGQVKYK 52
Db 19 WSPFGRRAIGRSRDCDWQIDDDNERRVSKLHCTLSRDGEGFIILDSANGTLVDGRLLLE 78
Qy 53 NSNQLSQSGDEITVG 67
Db 79 GESARLRDGSQINIG 93

RESULT 8
G98197
hypothetical protein AGR_L_1057 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: G98197
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G98197
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <KUR>
A:Cross-references: UNIPROT:Q8U7W4; GB:AE007870; PIDN:AAK89105.1; PID:g15158909; GSPDB:G
C:Genetics:
A:Gene: AGR_L_1057
A:Map position: linear chromosome

Query Match 23.8%; Score 80; DB 2; Length 399;
Best Local Similarity 26.7%; Pred. No. 0.022;
Matches 20; Conservative 16; Mismatches 31; Indels 8; Gaps 2;

Qy 1 WTF-----GRNPACDYHL-GNISRLSNKHFOILXXXXXLLNDISTNGTWLNGQVKYK 52
Db 19 WSPFGRRAIGRSRDCDWQIDDDNERRVSKLHCTLSRDGEGFIILDSANGTLVDGRLLLE 78
Qy 53 NSNQLSQSGDEITVG 67
Db 79 GESARLRDGSQINIG 93

RESULT 9
D70721
hypothetical protein Rv1827 - Mycobacterium tuberculosis (strain H37Rv)
```

C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: D70721
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70721
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-162 <COL>
A;Cross-references: UNIPROT:Q50606; GB:Z78020; GB:AL123456; NID:g3261625; PIDN:CAB01474.
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1827

Query Match 23.7%; Score 79.5; DB 2; Length 162;
Best Local Similarity 23.8%; Pred. No. 0.0095;
Matches 21; Conservative 16; Mismatches 15; Indels 21; Gaps 4;
Qy 4 GRNPACDYHLGNIS-----RLSNKHQFQLXXXXXXLLNDI-STNGTWLNGQVKVKNLS 54
Db 80 GRHPDSDIFLDVTVSRRAHAEFRLENNEFNVV-----DVGSLNGTVVNPVD--- 127
Qy 55 NQLLSQGDITVG 67
Db 128 SAVLANGDEVQIG 140

RESULT 10
A96950
DNA segregation ATP-ase FtsK/SpoIIIE (three ATPases), contains FHA domain [imported] - C
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: A96950
R;Nolling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A96950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1524 <KUR>
A;Cross-references: UNIPROT:Q97LZ2; GB:AE001437; PIDN:AAK78388.1; PID:g15023260; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0408

Query Match 23.4%; Score 78.5; DB 2; Length 1524;
Best Local Similarity 32.8%; Pred. No. 0.15;
Matches 22; Conservative 15; Mismatches 25; Indels 5; Gaps 3;
Qy 1 WTFGRNACDYHLGNISRLSNKHQFQLXXXXXXLLNDI-STNGTWLNGQVKVKNLS 59
Db 115 FTIGRGFENDIVFDI-KVSEKHAIEFDNGKYVLVDLNTNKTYLNGEMI---TRAILS 170
Qy 60 QGDITV 66
Db 171 EDDQINI 177

RESULT 11
G86467
hypothetical protein F7p12.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86467
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86467
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1587 <STO>
A;Cross-references: UNIPROT:Q9C8N3; GB:AE005172; NID:g10092383; PIDN:AAG12790.1; GSPDB:B:G
C;Genetics:
A;Map position: 1

Query Match 23.2%; Score 78; DB 2; Length 1587;
Best Local Similarity 29.9%; Pred. No. 0.18;
Matches 20; Conservative 16; Mismatches 27; Indels 4; Gaps 3;
Qy 4 GRNPACDYHLGNISRLSNKHQFI--LXXXXXXLLNDIST-NGTWLNGQVKVKNLSQ 60
Db 177 GRHPDCDILLTHPS--ISRFHLEIRSISSRQKLPFTDLSSVHGIVWRDLRIEPHGCVEVEE 235
Qy 61 GDEITVG 67
Db 236 GDTIRIG 242

RESULT 12
AF2214
hypothetical protein alr3269 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF2214
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2214
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <KUR>
A;Cross-references: UNIPROT:Q8YS22; GB:BA000019; PIDN:BAB74968.1; PID:g17132364; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3269

Query Match 22.5%; Score 75.5; DB 2; Length 234;
Best Local Similarity 30.3%; Pred. No. 0.046;
Matches 23; Conservative 12; Mismatches 22; Indels 19; Gaps 4;
Qy 1 WTFGR---NPACDY-----HLGNISRLSNKHQFQLXXXXXXLLNDISTNGTWLNGQVKE 51
Db 91 WTIGRDRHNGICITYDKLSRHHAAIKYVENQGF-----LLIDFQSTGSPVNGEPVY 142
Qy 52 KNSNQLLSQGDITVG 67
Db 143 QPI--ILKGDGRVRLG 156

RESULT 13
T44758
hypothetical protein MLCB1788.36c [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44758
R;Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z22833

A:Accession: T44758
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-162 <PAR>
A:Cross-references: UNIPROT:O32919; EMBL:AL008609; PIDN:CAA15468.1
A:Experimental source: cosmid B1788
C:Genetics:
A:Note: MLCB1788.36C

Query Match 22.2%; Score 74.5; DB 2; Length 162;
Best Local Similarity 29.2%; Pred. No. 0.042;
Matches 19; Conservative 23; Mismatches 23; Indels 5; Gaps 3;
Db 80 GRHPDSIFLDDVT-VSRRAEFLGNEFHVDVGSINGTYVNRFPD---SAVLANGD 135
Qy 63 EITVG 67
136 EVQIG 140

RESULT 14
S43941
N:Alternate names: protein D2370; protein YDL101c
C:Species: Saccharomyces cerevisiae
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43941; S67643; S67418; S72106
R:Zhou, Z.; Ellledge, S.J.
Cell 75, 1119-1127, 1993
A:Title: DUN1 encodes a protein kinase that controls the DNA damage response in yeast.
A:Reference number: S43941; MUID:94084787; PMID:8261511
A:Accession: S43941
A:Molecule type: DNA
A:Residues: 1-513 <ZHO>
A:Cross-references: UNIPROT:P39009; EMBL:L25548; NID:g435616; PIDN:AAAL6324.1; PID:g4356
R:Ballasta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.; Bo
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67629
A:Accession: S67643
A:Molecule type: DNA
A:Residues: 1-513 <BAL>
A:Cross-references: EMBL:Z74149; NID:g1431139; PIDN:CAA98668.1; PID:g1431140; GSPDB:GN00
A:Experimental source: strain S288C
R:Bozkovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jimine
submitted to the EMBL Data Library, February 1996
A:Reference number: S67406
A:Accession: S67418
A:Molecule type: DNA
A:Residues: 1-513 <BOS>
A:Cross-references: EMBL:X95644; NID:g1199535; PIDN:CRA64912.1; PID:g1199548
R:Saiz, J.E.; Buitrage, M.J.; Garcia, R.; Revuelta, J.L.; del Rey, F.
Yeast 12, 1077-1084, 1996
A:Title: The sequence of a 20.3 kb DNA fragment from the left arm of Saccharomyces cerev
A:Reference number: S72094; MUID:97051597; PMID:8896274
A:Accession: S72106
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-513 <SAI>
A:Cross-references: EMBL:X95644; NID:g1199535; PIDN:CAA64912.1; PID:g1199548
C:Genetics:
A:Gene: SGD:DUN1; MIPS:YDL101C
A:Cross-references: MIPS:YDL101c; SGD:S0002259
A:Map position: 4L
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: protein kinase DUN1; kinase interaction domain homology; protein kinase H
C:Keywords: ATP; nucleus; phosphoprotein; phosphotransferase; serine/threonine-specific
F:58-128/Domain: kinase interaction domain homology <KIH>
F:198-480/Domain: protein kinase homology <KIN>
F:206-214/Region: protein kinase ATP-binding motif

Query Match 22.0%; Score 74; DB 1; Length 513;
Best Local Similarity 29.1%; Pred. No. 0.17; Mismatches 14; Indels 20; Gaps 4;
Matches 23; Conservative 22

Qy 2 TFGRNPCADYHLG-----NISRLNKHQFQILXXXXXXLLN--DISTNGTTLNGQ 48
Db 57 TIGRSRSCVILSEPDISTFHAERHLLQMDVDFQ-----RNLIINIDKSRNGTFINGN 110

Qy 49 KVEKNSNQLSQGDEITVG 67
Db 111 RLVK-KDYILKNGDRIVFG 128

RESULT 15

B71407

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: columbia

C>Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004

C:Accession: B71407

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
avanagh, T.; 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalmatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: B71407

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-386 <BEV>

A:Cross-references: UNIPROT:O23305; GB:Z97336; NID:g2244788; PID:e326894; PID:g2244805

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match

21.3%; Score 71.5; DB 2; Length 386;

Best Local Similarity 27.8%; Pred. No. 0.26;

Matches 20; Conservative 12; Mismatches 29; Indels 11; Gaps 2;

Qy 7 PACDYHLGNISR-----LSNKHQFQILXXXXXXLLNDI-STNGTTLNGQKVEKNSN 55

Db 24 PGSTIRGVRVIRGNEIAKAGISTKHLRIEDSGNWWIQDGLSSNGTLLNSALDPETS 83

Qy 56 QLLSQGDEITVG 67

Db 84 VNLGDDGVIKLG 95

RESULT 16

T02045

kinase associated protein phosphatase - maize

C:Species: Zea mays (maize)

C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004

C:Accession: T02045

R:Braun, D.M.; Stone, J.M.; Walker, J.C.

A:Submitted to the EMBL Data Library, December 1996

A:Description: Multiple receptor-like protein kinases interact with the KI domain of the
es.

A:Reference number: Z14510

A:Accession: T02045

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-583 <BEA>

A:Cross-references: UNIPROT:O49973; EMBL:U81960; NID:g2735007; PIDN:AAB93832.1; PID:g273

A:Experimental source: strain B73

C:Genetics:

A:Gene: KAPP

C:Superfamily: kinase associated protein phosphatase; kinase interaction domain homology
F:204-278/Domain: kinase interaction domain homology <KIH>

Query Match 20.8%; Score 70; DB 2; Length 583;
Best Local Similarity 36.4%; Pred. No. 0.64;
Matches 28; Conservative 7; Mismatches 30; Indels 12; Gaps 3;

QY 2 TFCGRNPACDYHLGNISRLSNKHFOIL---XXXXXXLLNDISTNGTWLNGQVKEKNS-----54
DB 203 TLGRVPSPDLVLKD-SEVSGKHAQINWNGKTLKWLVDVMSLNGTFLNSQAVHPSAGSR 261

QY 55 ----NQLLSQDDEITVG 67
DB 262 HWGEPAPLNGDITLG 278

RESULT 17
S76488
hypothetical protein - Synchocystis sp. (strain PCC 6803)
C;Species: Synchocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76488
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M. Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <KAN>
A;Cross-references: UNIPROT:P74513; EMBL:D90915; GB:AB001339; NID:G1653604; PIDN:BAA1861
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 20.7%; Score 69.5; DB 2; Length 157;
Best Local Similarity 28.6%; Pred. No. 0.18; Mismatches 15; Indels 13; Gaps 4;
Matches 22; Conservative 15

QY 1 WTEGRNPACDYHLGN---ISR---LSNKHFOILXXXXXXLLND-----ISTNGTWLNGQKV 50
DB 32 YTGSRPADIRIKSQFVSRHAHLVLRKSSDDVQAARIIDGDEDDGSSVNGLMINGKV 91

QY 51 EKNSNQLLSQDDEITVG 67
DB 92 QEH---IIQTGDEITVG 105

RESULT 18
AB2022
hypothetical protein alr1728 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB2022
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2022
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-445 <KUR>
A;Cross-references: UNIPROT:Q8YW88; GB:BA000019; PIDN:BAB73427.1; PID:gl7130818; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1728

Query Match 20.7%; Score 69.5; DB 2; Length 445;
Best Local Similarity 31.5%; Pred. No. 0.55; Mismatches 14; Indels 11; Gaps 4;
Matches 23; Conservative 14

QY 4 GRNPACDYHLGNI--SRLSNKHFOI-----LXXXXXXLLNDI--STNGTWLNGQVKEKNS 54
DB 394 GRGDAQFRLPD-TGVSRRLHLEIRWDGQVALLSGLNSTNGTNNAPVQE---WQLADGD 449

Db 38 GRDPSCQVVLDAMMYRMVSRHAWRPVASSVDKFSWVLCDLNSANGTYLNGQRL--YG 95
QY 55 NQLLSQDDEITVG 67
DB 96 CQELHAGDRISLG 108

RESULT 19
AE1899
adenylate cyclase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE1899
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1899
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-546 <KUR>
A;Cross-references: UNIPROT:P94184; GB:BA000019; PIDN:BAB72700.1; PID:gl7130088; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: cyad

Query Match 20.7%; Score 69.5; DB 2; Length 546;
Best Local Similarity 33.3%; Pred. No. 0.69; Mismatches 14; Indels 5; Gaps 5;
Matches 23; Conservative 14

QY 1 WTEGRNPACDYHLGNISRLSNKHFOILXXXXXXLLNDI--STNGTWLNGQVKEKNSQLL 58
DB 24 FTIGRLPECNLYL-PPAGVSRRAQLVKADGKWIIEDLGSKNGTQVN-QSIVSHPRQ-L 80

QY 59 SQDDEITVG 67
DB 81 QHGDVILWG 89

RESULT 20
T10015
hypothetical protein MLB1770.15c - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10015
R;Cole, S.T.
A;Reference number: Z16916
A;Accession: T10015
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-463 <COL>
A;Cross-references: UNIPROT:Q50190; EMBL:Z70722; NID:e1059634; PID:e3337961
C;Genetics:
A;Note: MLB1770.15c

Query Match 20.1%; Score 67.5; DB 2; Length 463;
Best Local Similarity 32.3%; Pred. No. 1; Mismatches 13; Indels 5; Gaps 3;
Matches 21; Conservative 13

QY 4 GRNPACDYHLGNISRLSNKHFOILXXXXXXLLNDI--STNGTWLNGQVKEKNSQLLSQGD 62
DB 394 GRGDAQFRLPD-TGVSRRLHLEIRWDGQVALLSGLNSTNGTNNAPVQE---WQLADGD 449

QY 63 EITVG 67
DB 450 VIRLG 454

RESULT 21
P86911

conserved hypothetical protein ML0022 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F86911
R:; Cole, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; HC
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; SC
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: F86911
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-488 <STO>
A:Cross-references: UNIPROT:Q9CDE4; GB:AL450380; NID:gl3092432; PIDN:CAC29530.1; GSPDB:C
C:Genetics:
A:Gene: ML0022

Query Match 20.1%; Score 67.5; DB 2; Length 488;
Best Local Similarity 32.3%; Pred. No. 1.1;
Matches 21; Conservative 13; Mismatches 26; Indels 5; Gaps 3;

Qy 4 GRNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWNKGQKVKNSNQLLSQGD 62
Db 419 GRGQDAQFLRPD-TGVSRRLHLEIRWDGQVALLSDLNSTNGTNNAPVQE---WQLADGD 474

Qy 63 EITVG 67
Db 475 VIRLG 479

RESULT 22
B48666
cell proliferation antigen Ki-67, short form - human
C:Species: Homo sapiens (man)
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 15-Mar-2004
C:Accession: B48666
R:Schlueter, C.; Duchrow, M.; Wohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Gerde
J. Cell Biol. 123, 513-522, 1993
A:Title: The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiq
ins.
A:Reference number: A48666; MUID:94043435; PMID:8227122
A:Accession: B48666
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2897 <SCH>
A:Cross-references: EMBL:X65551
C:Keywords: alternative splicing; cell cycle control; nucleus; tandem repeat
F:29-91/Domain: kinase interaction domain homology <KIH>

Query Match 19.9%; Score 67; DB 2; Length 2897;
Best Local Similarity 30.8%; Pred. No. 8.9;
Matches 20; Conservative 13; Mismatches 28; Indels 4; Gaps 3;

Qy 3 FGRNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWNKGQKVKNSNQLLSQG 61
Db 29 FGRGIECDIRI-QLPVPVSKQHCKIEIHEQEAILHNFSTNPTQVNGSVIDEFVR--LKHG 85

Qy 62 DEITV 66
Db 86 DVIIT 90

RESULT 23
T30249
cell proliferation antigen Ki-67 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30249
R:Starborg, M.; Gell, K.; Brundell, E.; Hoog, C.
J. Cell Sci. 109, 143-153, 1996
A:Title: The murine Ki-67 cell proliferation antigen accumulates in the nucleolar and he

1 for cell cycle progression.
A:Reference number: 220787; MUID:96431717; PMID:8834799
A:Accession: T30249
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2938 <STA>
A:Cross-references: UNIPROT:Q61769; EMBL:X82786; NID:gl177527; PIDN:CAA58026.1; PID:gl17
A:Experimental source: strain CBA; testis
C:Genetics:
A:Gene: Ki-67
C:Keywords: cell cycle control; nucleus; tandem repeat

Query Match 19.9%; Score 67; DB 2; Length 2938;
Best Local Similarity 32.3%; Pred. No. 9;
Matches 21; Conservative 12; Mismatches 28; Indels 4; Gaps 3;

Qy 3 FGRNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWNKGQKVKNSNQLLSQG 61
Db 29 FGRGIECDIRI-QLPVPVSKQHCKIEIHEQEAILHNFSTNPTQVNGSVIDEFVR--LRHG 85

Qy 62 DEITV 66
Db 86 DVIIT 90

RESULT 24
A48666
cell proliferation antigen Ki-67, long form - human
C:Species: Homo sapiens (man)
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: A48666
R:Schlueter, C.; Duchrow, M.; Wohlenberg, C.; Becker, M.H.G.; Key, G.; Flad, H.D.; Gerde
J. Cell Biol. 123, 513-522, 1993
A:Title: The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiq
ins.
A:Reference number: A48666; MUID:94043435; PMID:8227122
A:Accession: A48666
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-3256 <SCH>
A:Cross-references: UNIPROT:P46013; EMBL:X65550; NID:gl15818; PIDN:CAA46519.1; PID:gl158
C:Keywords: alternative splicing; cell cycle control; nucleus; tandem repeat
F:29-91/Domain: kinase interaction domain homology <KIH>

Query Match 19.9%; Score 67; DB 2; Length 3256;
Best Local Similarity 30.8%; Pred. No. 10;
Matches 20; Conservative 13; Mismatches 28; Indels 4; Gaps 3;

Qy 3 FGRNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDI-STNGTWNKGQKVKNSNQLLSQG 61
Db 29 FGRGIECDIRI-QLPVPVSKQHCKIEIHEQEAILHNFSTNPTQVNGSVIDEFVR--LKHG 85

Qy 62 DEITV 66
Db 86 DVIIT 90

RESULT 25
A83872
hypothetical protein BH1777 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: A83872
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A83872
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <STO>
A:Cross-references: UNIPROT:Q9KBZ7; GB:BA001513; GB:BA000004; NID:gl0174345; PIDN:BA8054
A:Experimental source: strain C-125

[illegible]

Db 514 VIRLG 518 | : |

RESULT 30

E86579

FHA domain (imported) - Chlamydomophila pneumoniae (strain J138)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E86579

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: E86579

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-845 <STO>

A:Cross-references: UNIPROT:Q927J3; GB:BA000008; NID:98979084; PIDN:BAA98919.1; GSPDB:GN

A:Experimental source: strain J138

C:Genetics:

A:Gene: CPJ0712

Query Match 18.8%; Score 63; DB 2; Length 845;

Best Local Similarity 29.4%; Pred. No. 7.6;

Matches 20; Conservative 19; Mismatches 21; Indels 8; Gaps 5;

Qy 1 WTFGRNP-ACDYHLGNISRLSKHFQI-LXXXXXXLLNDI-STNGTWLNGQKVEK-----N 53

Db 422 YILGDTPTTCDIVPNDLS-VSHQAKITVNGDGGILIEDLDSKNGVIGRKGIDKIDTSTLS 480

Qy 54 SNQLLSOG 61

Db 481 SNQVVALG 488

RESULT 31

H72046

Fha domain (homolog to adenylate cyclase) - Chlamydomophila pneumoniae (strains CWL029 and

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: H72046; A81620

R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: H72046

A:Molecule type: DNA

A:Residues: 1-845 <REN>

A:Cross-references: UNIPROT:Q927J3; GB:AE001652; NID:94376997; PIDN:AAD1885

A:Experimental source: strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: A81620

A:Molecule type: DNA

A:Residues: 1-845 <REA>

A:Cross-references: GB:AE002167; GB:AE002161; NID:97186971; PIDN:AAF37929.1; PID:9718897

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CPN0712; CP0034

Query Match 18.8%; Score 63; DB 2; Length 845;

Best Local Similarity 29.4%; Pred. No. 7.6;

Matches 20; Conservative 19; Mismatches 21; Indels 8; Gaps 5;

Qy 1 WTFGRNP-ACDYHLGNISRLSKHFQI-LXXXXXXLLNDI-STNGTWLNGQKVEK-----N 53

Db 422 YILGDTPTTCDIVPNDLS-VSHQAKITVNGDGGILIEDLDSKNGVIGRKGIDKIDTSTLS 480

Qy 54 SNQLLSOG 61

Db 481 SNQVVALG 488 | : : : |

RESULT 32

T00393

Nijmegen breakage syndrome protein NBS1 - human

N:Alternate names: cell cycle regulatory protein p95

C:Species: Homo sapiens (man)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C:Accession: T00393

R:Matsumura, S.; Tauchi, H.; Nakamura, A.; Kondo, N.; Sakamoto, S.; Endo, S.; Smeets, D.;

Nature Genet. 19, 179-181, 1998

A:Title: Positional cloning of the gene for Nijmegen breakage syndrome.

A:Reference number: Z14144; MUID:98282099; PMID:9620777

A:Accession: T00393

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-754 <MAT>

A:Cross-references: UNIPROT:O60934; EMBL:AB013139; PIDN:BAA28616.1

C:Genetics:

A:Gene: NBS1

A:Map position: 8q21

A:Introns: 13/1; 57/3; 107/2; 160/3; 195/2; 234/3; 299/2; 332/1; 375/2; 466/2; 615/3; 63

C:Superfamily: human Nijmegen breakage syndrome protein NBS1

C:Keywords: DNA repair

Query Match 18.5%; Score 62; DB 2; Length 754;

Best Local Similarity 39.4%; Pred. No. 9.1;

Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 35 LNDISTNGTWLNGQKVEKSNQLLSQGDDEITVG 67

Db 68 LKNSKYGTFVNBKQMONGFSRTLKSGDGITFG 100

RESULT 33

D70986

Probable ABC transporter Rv1747 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: D70986

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Bayrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70986

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-865 <COL>

A:Cross-references: UNIPROT:O65934; GB:Z95890; GB:AL123456; NID:g3242245; PIDN:CAB09333.

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1747

C:Superfamily: Mycobacterium tuberculosis probable ABC transporter Rv1747; ATP-binding c

C:Keywords: ATP

F:335-528/Domain: ATP-binding cassette homology <ABC>

Query Match 18.3%; Score 61.5; DB 1; Length 865;

Best Local Similarity 44.8%; Pred. No. 12;

Matches 13; Conservative 6; Mismatches 7; Indels 3; Gaps 1;

Qy 39 STNGTWLNGQKVEKSNQLLSQGDDEITVG 67

Db 268 SINGTFVNGARVDA---ALLHGDGVITIG 293

RESULT 34

T49879

hypothetical protein T211.110 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49879
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T49879
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1075 <REV>
A;Cross-references: UNIPROT:O9LY20; EMBL:AL163912; GSPDB:GN00063; ATSP:T211.110
A;Experimental source: cultivar Columbia; BAC clone T211
C;Genetics:
A;Gene: ATSP:T211.110
A;Map position: 5
A;Introns: 411/2; 456/2; 504/3; 604/3; 685/1; 731/3; 768/3; 786/1; 822/3; 865/3; 895/3
Query Match 18.3%; Score 61.5; DB 2; Length 1075;
Best Local Similarity 23.5%; Pred. No. 15;
Matches 24; Conservative 14; Mismatches 25; Indels 39; Gaps 4;
QY 1 WTGRNPA---CDYHLCNLSLNKHFQILXXXXXIL----- 35
Db 40 YTGSSSDGFCDFVIDH-SISRRKHCQILFDSQSHKLYIFDGVHLPSPGFSQVDEFR 98
QY 36 -----NDI-----STNGTWLNGQVKVKNLNQLSQGDEI 64
Db 99 RRLVGVEDLGNLFRASLNGVYVNRVRKSKQVEVSDIDVE 140
RESULT 35
D96904
probable membrane protein, containing FHA domain [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: D96904
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D96904
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <KUR>
A;Cross-references: UNIPROT:Q97N03; GB:A5001437; PIDN:AAK78023.1; PID:gl5022857; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0036
Query Match 18.0%; Score 60.5; DB 2; Length 468;
Best Local Similarity 31.3%; Pred. No. 8.4;
Matches 21; Conservative 11; Mismatches 20; Indels 15; Gaps 4;
QY 11 YHLGNIS-----RLSNKHQFQILXXXXXLLNDI-STNGTWLNGQVKVKNLNQL-- 57
Db 392 FKIGRISQADYISDNKAVGLHAIRKQNEKYVILDLTSRNGTFVNGQXI--NSDELXE 449
QY 58 LSQGDEI 64
Db 450 IRNGDTI 456
RESULT 36
T19167
hypothetical protein C09H10.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19167
R;Ainscough, R.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19082
A;Accession: T19167

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-150 <WIL>
A;Cross-references: UNIPROT:Q3U302; EMBL:Z50109; PIDN:CAB54192.1; GSPDB:GN00020; CESP:CO
A;Experimental source: Clone C09H10
C;Genetics:
A;Gene: CESP:C09H10.10
A;Map position: 2
A;Introns: 15/1; 58/3
C;Superfamily: Caenorhabditis elegans hypothetical protein C09H10.10
Query Match 17.9%; Score 60; DB 2; Length 150;
Best Local Similarity 27.1%; Pred. No. 2.8;
Matches 19; Conservative 12; Mismatches 31; Indels 8; Gaps 3;
QY 3 FGR-NPAC----DYHLGNTSLSLNKHFQILXXXXXLLNDISTNGTWLNGQVKVKNLNQL 57
Db 28 FGREKKVCHITFDPHAARVSR-----HASIEWGDEGLFTDKSKEGTEINGTRLKQSSQEL 84
QY 58 LSQGDEITVG 67
Db 85 HEGVQLAIG 94
RESULT 37
T16191
hypothetical protein F27D9.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16191
R;Bentley, D.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid F27D9.
A;Reference number: Z18473
A;Accession: T16191
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-431 <BEN>
A;Cross-references: UNIPROT:Q19846; EMBL:U49829; NID:gl203924; PID:gl203931; PIDN:AAA933
C;Genetics:
A;Gene: CESP:F27D9.8
A;Introns: 8/1; 52/3; 74/1; 96/1; 128/2; 149/3; 199/2; 255/3; 297/3; 319/3; 350/3; 390/3
Query Match 17.9%; Score 60; DB 2; Length 431;
Best Local Similarity 36.7%; Pred. No. 8.9;
Matches 11; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
QY 38 ISTNGTWLNGQVKVKNLNQLSQGDEITVG 67
Db 87 VEVNGISIEGQSHDEVVNMKSSGDQVTLG 116
RESULT 38
S51457
probable membrane protein YLR238w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8083.1
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S51457
R;Hallsworth, K.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmid 8083.
A;Reference number: S51443
A;Accession: S51457
A;Molecule type: DNA
A;Residues: 1-478 <HAL>
A;Cross-references: UNIPROT:Q06001; EMBL:U19027; NID:g609363; PID:g609364; GSPDB:GN00012.
C;Genetics:
A;Gene: MIPS:YLR238w
A;Cross-references: SGD:S0004228
A;Map position: 12R
C;Keywords: transmembrane protein
F;455-471/Domain: transmembrane #status predicted <TM>

Query Match 17.9%; Score 60; DB 2; Length 478;
Best Local Similarity 30.5%; Pred. No. 10;
Matches 18; Conservative 15; Mismatches 20; Indels 6; Gaps 4;

Qy 14 GNI-SRLSNKHFIQLX---XXXXXLLNDI-STNGTWTWNGQVKVKNQLLSQGDEITVG 67
Db 135 GNFDRLVSRNHALSCDPLTGKVIYRDLKSSNOTFINGQIGSNQDVE-IKVGDIVDLG 192

RESULT 39

S75331
penicillin-binding protein 1B mrcB - *Synechocystis* sp. (strain PCC 6803)
N;Alternate names: protein slr1710
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75331
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75331
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-749 <KAN>
A;Cross-references: UNIPROT:P73218; EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAAL1724
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: mrcB
C;Superfamily: penicillin-binding protein

Query Match 17.9%; Score 60; DB 2; Length 749;
Best Local Similarity 27.9%; Pred. No. 16;
Matches 19; Conservative 15; Mismatches 26; Indels 8; Gaps 3;

Qy 4 GRNPACDYHLGNISRLSNKHFIQLX-----XXXXXLLNDISTNGTWTWNGQVKVKNQLLS 59
Db 64 GRSSCDIQLQN-PLVSGTCHSLRRDPDPHPNQFFIRDEGSSNGIYLQRRRLK---SYRLQ 119

Qy 60 QGDEITVG 67

Db 120 HGDEITLG 127

RESULT 40

D88188
protein C18H9.1 [imported] - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D88188
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D88188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-192 <STO>
A;Cross-references: UNIPROT:Q09481; GB:chr II; PIDN:C46686.1; PID:g722381; GSPDB:GN00020
A;Note: similar to *E. faecium* initiation factor (IF-2)
C;Genetics:
A;Gene: C18H9.1
A;Map position: 2

Query Match 17.7%; Score 59.5; DB 2; Length 192;
Best Local Similarity 26.3%; Pred. No. 4.3;
Matches 15; Conservative 7; Mismatches 14; Indels 21; Gaps 1;

Qy 5 RNPACDYHLGNISRLSNKHFIQLXXXXXXLLNDISTNGTWTWNGQVKVKNQLLSQG 61
Db 95 KNNRSNVHIGNIGNI-----CGNIGNCGKMMNRKNKNGGRNQG 130

Search completed: February 24, 2005, 15:24:30
Job time : 30.1463 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 15:04:15 ; Search time 125.285 Seconds
(without alignments)
273.851 Million cell updates/sec

Title: US-09-936-956-19

Perfect score: 336

Sequence: 1 WTGRPNPACDYLHGNISRLS.....QKVEKNSQLLSQGDIEITVG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	313.5	93.3	821	1 RA53_YEAST	P22216 saccharomyc
2	313.5	93.3	821	2 Q8B1S1	Q8B1S1 saccharomyc
3	273.5	81.4	767	2 Q6FK01	Q6fk01 candida gla
4	234	69.6	837	2 Q75CE9	Q75ce9 ashbya goss
5	233.5	69.5	813	2 Q6CKR2	Q6ckf2 kluyveromyc
6	171	50.9	751	2 Q6BUB2	Q6hub2 debaryomyc
7	137.5	40.9	204	2 Q68329	Q68329 schizosacch
8	137.5	40.9	460	1 CDS1_SCHPO	Q99170 schizosacch
9	116	34.5	503	2 Q90ZV5	Q90zy5 brachydanio
10	115	34.2	898	2 Q6C079	Q6c079 yarrowia li
11	112	33.3	503	2 Q803B2	Q803e2 brachydanio
12	105.5	31.1	504	2 Q87HC2	Q87pu0 vibrio para
13	104.5	31.1	398	2 Q87HC2	Q87hc2 vibrio para
14	100.5	29.9	359	2 Q98IK6	Q98ik6 rhizobium l
15	99	29.5	626	2 Q6C079	Q6c079 yarrowia li
16	98.5	29.3	460	2 Q89P85	Q89p85 bradyrhizob
17	95	28.3	1029	2 Q8DJ80	Q8dj88 synecococc
18	93.5	27.8	567	2 Q9FWA0	Q9fwa0 arabidopsis
19	93.5	27.8	585	2 Q9M8A0	Q9m8a0 arabidopsis
20	92.5	27.5	443	2 Q6BH59	Q6bh59 debaryomyc
21	91	27.1	374	2 Q8D6T0	Q8d6t0 vibrio vuln
22	91	27.1	434	2 Q7MDP1	Q7mdp1 vibrio vuln
23	89.5	26.6	712	2 Q7SEK0	Q7seko neurospora
24	89	26.5	398	2 Q761W8	Q761w8 entameba h
25	87.5	26.0	436	2 Q6LUB6	Q6lue6 photobacter
26	87	25.9	159	2 Q97LQ0	Q97lq0 clostridium
27	86.5	25.7	445	1 MEK1_SCHPO	Q10292 schizosacch
28	85.5	25.4	497	2 Q91751	Q91751 pseudomonas
29	85	25.3	664	1 CHER_HUMAN	Q96ep1 homo sapien
30	85	25.3	674	1 Q8MQJ2	Q8mqj2 bdellovibri
31	84	25.0	342	2 Q7UIF2	Q7uif2 rhodospirill

RESULT 1

RA53_YEAST	STANDARD;	PRT;	821 AA.
ID	RA53_YEAST	STANDARD;	PRT;
AC	P22216;		
DT	01-AUG-1991 (Rel. 19, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	25-OCT-2004 (Rel. 45, Last annotation update)		
DE	Serine/threonine-protein kinase RAD53 (EC 2.7.1.37) (Serine-protein kinase 1).		
DE	kinase 1).		
GN	Name=RAD53; Synonyms=MBC2, SAD1, SPK1; OrderedLocusNames=YPL153C;		
GN	ORFNames=P2588;		
OS	Saccharomycetes cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91117267; PubMed=1899289;		
RA	Stern D.F., Zheng P., Beidler D.R., Zerillo C.;		
RT	"Spk1, a new kinase from Saccharomycetes cerevisiae, phosphorylates proteins on serine, threonine, and tyrosine.";		
RL	Mol. Cell. Biol. 11:987-1001(1991).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288C / AB972;		
RX	MEDLINE=97103777; PubMed=8948103;		
RX	DOI=10.1002/(SICI)1097-0061(199611)12:14<1483::AID-YEA34>3.3.CO;2-F;		
RA	Purnelle B., Coster F., Goffeau A.;		
RT	"The sequence of 55 kb on the left arm of yeast chromosome XVI identifies a small nuclear RNA, a new putative protein kinase and two new putative regulators.";		
RT	Yeast 12:1483-1492(1996).		
RL	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288C / AB972;		
RX	MEDLINE=97313271; PubMed=9169875;		
RA	Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W., Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V., Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E., Churcher C.M., Coster F., Davis K., Davis R.W., Dietrich F.S., Dilius H., DiPaolo T., Dubois E., Duesterhoeft A., Duncan M., Floeth U., Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W., Hunkeler-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K., Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D., Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D., Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M., Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H., Urrestarazu L.A., Ushinsky S., Vierendeels F., Viessers S., Voss H., Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E., Zhong W.W., Zollner A., Vo D.H., Hani J.;		
RT	"The nucleotide sequence of Saccharomycetes cerevisiae chromosome XVI.";		
RL	Nature 387:103-105(1997).		

ALIGNMENTS

32	84	25.0	664	1	CHFR_MOUSE
33	83.5	24.9	143	2	Q6NHD4
34	83.5	24.9	288	2	Q6NKG7
35	83.5	24.9	298	2	Q8FUH9
36	82.5	24.6	557	2	Q6MQJ4
37	81.5	24.3	142	2	Q8FTJ5
38	81.5	24.3	288	2	Q8NU92
39	81.5	24.3	298	2	Q6MBV2
40	81	24.1	952	1	YK15_CAEEL
41	80	23.8	399	2	Q8U7W4
42	80	23.8	743	1	YK18_SCHPO
43	79.5	23.7	159	2	Q73ZB1
44	79.5	23.7	162	1	YI27_MYCTU
45	79.5	23.7	162	1	YI58_MYCBO

Q81013	mus musculus
Q8nhd4	corynebacte
Q8nkg7	corynebacte
Q8fuh9	corynebacte
Q6mqj4	bdellovibri
Q8ftj5	corynebacte
Q8nu92	corynebacte
Q6mbv2	corynebacte
P46012	caenorhabdi
Q8u7w4	agrobacteri
O14270	schizosacch
Q73zr1	mycobacteri
P64897	mycobacteri
P64898	mycobacteri

RA RP MEDLINE=95047382; PubMed=7958905;
RA Allen J.B., Zhou Z., Siede W., Friedberg E.C., Elledge S.J.;
RT "The SADI/RAD53 protein kinase controls multiple checkpoints and DNA
RL damage-induced transcription in yeast.";
RL Genes Dev. 8:2401-2415(1994).
RA [5]
RA MEDLINE=20018334; PubMed=10550056; DOI=10.1126/science.286.5442.1166;
RA Sanchez Y., Bachant J., Wang H., Hu F., Liu D., Fetzlaff M.,
RA Elledge S.J.;
RT "Control of the DNA damage checkpoint by chk1 and rad53 protein
RT kinases through distinct mechanisms.";
RL Science 286:1166-1171(1999).
RA [6]
RA PHOSPHORYLATION.
RX MEDLINE=20031667; PubMed=10562568; DOI=10.1093/emboj/18.22.6561;
RA Pelliccioli A., Lucca C., Liberi G., Marini F., Lopes M., Plevani P.,
RA Romano A., Di Fiore P.P., Foiani M.;
RT "Activation of Rad53 kinase in response to DNA damage and its effect
RT in modulating phosphorylation of the lagging strand DNA polymerase.";
RL EMBO J. 18:6561-6572(1999).
RA [7]
RA STRUCTURE BY NMR OF 573-730.
RX PubMed=11809875; DOI=10.1093/nar/30.3.643;
RA Kim E.M., Jang Y.K., Park S.D.;
RT "Phosphorylation of Rph1, a damage-responsive repressor of PHR1 in
RT Saccharomyces cerevisiae, is dependent upon Rad53 kinase.";
RL Nucleic Acids Res. 30:643-648(2002).
CC -!- FUNCTION: Controls S-phase checkpoint as well as G1 and G2 DNA
CC damage checkpoints. Phosphorylates proteins on serine, threonine,
CC and tyrosine. Prevents entry into anaphase and mitotic exit after
CC DNA damage via regulation of the Polo kinase CDC5. Seems to be
CC involved in the phosphorylation of RPH1.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -!- SIMILARITY: Contains 2 FHA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55623; AAA35070.1; -;
DR EMBL; X98770; CAA65568.1; -;
DR PIR; A39616; A39616.
DR PDB; 1DMZ; NMR; A=573-730.
DR PDB; 1PHQ; NMR; A=573-730.
DR PDB; 1FHR; NMR; A=573-730.
DR PDB; 1G3G; NMR; A=1-164.
DR PDB; 1G6G; X-ray; A/B=29-155.
DR PDB; 1J4K; NMR; A=573-730.
DR PDB; 1J4L; NMR; A=573-730.
DR PDB; 1J4O; NMR; A=14-164.
DR PDB; 1J4P; NMR; A=14-164.
DR PDB; 1J4Q; NMR; A=14-164.
DR PDB; 1K2M; NMR; A=573-730.
DR PDB; 1K2N; NMR; A=573-730.
DR PDB; 1K3J; NMR; A=14-164.

DR PDB; 1K3N; NMR; A=14-164.
DR PDB; 1K3Q; NMR; A=14-164.
DR PDB; 1QU5; NMR; A=549-730.
DR GerMOnline; 144135; -;
DR SGD; S00006074; RAD53.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0004712; F:protein threonine/tyrosine kinase activity; IDA.
DR GO; GO:0006281; P:DNA repair; IMP.
DR GO; GO:0006139; P:nucleobase, nucleoside, nucleotide and nucl. . .; IGI.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 2.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 2.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW 3D-structure; ATP-binding; Cell cycle; DNA damage; Nuclear protein;
KW Phosphorylation; Repeat; Serine/threonine-protein kinase; Transferase;
KW Tyrosine-protein kinase.
FT DOMAIN 66 116 FHA 1.
FT DOMAIN 198 466 Protein kinase.
FT DOMAIN 601 664 FHA 2.
FT NP_BIND 204 212 ATP (By similarity).
FT BINDING 227 227 ATP (By similarity).
FT ACT_SITE 319 319 Proton acceptor.
FT STRAND 34 38
FT TURN 39 40
FT STRAND 46 48
FT HELIX 52 57
FT STRAND 62 69
FT TURN 72 73
FT STRAND 76 77
FT TURN 82 83
FT STRAND 89 94
FT TURN 95 96
FT STRAND 99 103
FT STRAND 109 111
FT TURN 112 113
FT STRAND 114 115
FT TURN 118 119
FT STRAND 122 123
FT TURN 126 127
FT STRAND 129 132
FT TURN 134 135
FT HELIX 137 139
FT STRAND 141 147
FT HELIX 149 153
FT TURN 154 154
FT STRAND 154 154
FT STRAND 578 582
FT TURN 584 585
FT STRAND 592 594
FT TURN 596 597
FT STRAND 601 604
FT TURN 607 608
FT STRAND 611 612
FT TURN 616 617
FT TURN 620 621
FT STRAND 623 630
FT STRAND 644 651
FT TURN 654 655
FT STRAND 657 659
FT TURN 660 661
FT STRAND 662 664
FT TURN 666 667
FT STRAND 668 671

QY	TURN	674	675
FT	STRAND	676	683
 Query Match 93.3%; Score 313.5; DB 1; Length 821; Best Local Similarity 89.7%; Pred. No. 7e-32; Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;			
QY	1	WTGFRNPACDYHLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQNLLS	59
Db	66	WTGFRNPACDYHLGNISRLSNKHFIQLLGEDGNLLNDISTNGTWLNGQVKVKNQNLLS	125
QY	60	QGDEITVG 67	
Db	126	QGDEITVG 133	
 RESULT 2 Q6BLS1 PRELIMINARY; PRT; 821 AA.			
ID	Q6BLS1	PRELIMINARY; PRT; 821 AA.	
AC	Q6BLS1	PRELIMINARY; PRT; 821 AA.	
DT	25-OCT-2004	(TEMBLrel. 28, Created)	
DT	25-OCT-2004	(TEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TEMBLrel. 28, Last annotation update)	
DE	VPL153C.		
GN	Name=RAU53;		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OX	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
NCBI_TaxID=4932;			
RN	[1]		
SEQUENCE FROM N.A.			
Marsischky G., Rolfs A., Richardson A., Kane M., Baqui M., Taycher E., Hu Y., Vannberg J., Weger J., Kramer J., Moreira D., Kelley F., Zuo D., Raphael J., Hogle C., Jepson D., Williamson J., Camargo A., Gonzaga L., Vasconcelos A.T., Simpson A., Kolodner R., Harlow E., LaBaer J.;			
"Creation of the YFLX clone resource: cloning of Saccharomyces cerevisiae ORFs in the Gateway recombinational cloning system."; RT Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.			
-!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.			
EMBL; AV693009; AT93028.1; --			
SGD; S00006074; RAD53.			
GO; GO:0005524; F:ATP binding; IEA.			
GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.			
GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.			
GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
InterPro; IPR000253; FHA.			
InterPro; IPR011009; Kinase like.			
InterPro; IPR000719; Prot_kinase.			
InterPro; IPR002230; Ser_thr_pkinase.			
InterPro; IPR008271; Ser_thr_pkin_AS.			
InterPro; IPR008984; SMAD_FHA.			
InterPro; IPR001245; Tyr_pkinase.			
Pfam; PF00498; FHA; 2.			
Pfam; PF00069; Pkinase; 1.			
ProDom; PD000001; Prot_kinase; 1.			
SMART; SM00240; FHA; 2.			
SMART; SM00220; S_TKc; 1.			
SMART; SM00219; Tyrc; 1.			
PROSITE; PS50006; FHA DOMAIN; 2.			
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.			
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.			
SEQUENCE 821 AA; 91934 MW; AAB353DC7DF68119 CRC64;			
Query Match 93.3%; Score 313.5; DB 2; Length 821; Best Local Similarity 89.7%; Pred. No. 7e-32; Matches 61; Conservative 0; Mismatches 6; Indels 1; Gaps 1;			
QY	1	WTGFRNPACDYHLGNISRLSNKHFI-LXXXXXXLLNDISTNGTWLNGQVKVKNQNLLS	59
Db	66	WTGFRNPACDYHLGNISRLSNKHFIQLLGEDGNLLNDISTNGTWLNGQVKVKNQNLLS	125

```
Db 125 QGDEITVG 132
|||||
RESULT 4
Q75CE9 PRELIMINARY; PRT; 837 AA.
AC Q75CE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ACRI142Wp.
GN ORFNames=ACRI142W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Brachat S., Voegel S.E., Dietrich F.S., Lerch A., Gaffney T.,
RA Philippsen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AE016816; AAS51368.1; -.
DR HSSP; P22216; 1K3Q.
DR AGD; ACRI142W; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00498; FHA; 2.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00006; FHA DOMAIN; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 837 AA; 93517 MW; DDB15B755394B627 CRC64;

Query Match 69.6%; Score 234; DB 2; Length 837;
Best Local Similarity 66.7%; Pred. No. 1.7e-21;
Matches 46; Conservative 6; Mismatches 15; Indels 2; Gaps 2;

Qy 1 WFGFRN-PACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWTWLNQGVKVNQLL 58
|||||
Db 93 WLFGRNQSCDYHLGDISRLSNKHFOIMLGEDGNLLKDLSTNGTWTWLNDRIEKGQNWIL 152
|||||
Qy 59 SQGDEITVG 67
|||||
Db 153 QGQDEITVG 161
|||||

RESULT 5
Q6CKF2 PRELIMINARY; PRT; 813 AA.
AC Q6CKF2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome F of strain NRRL Y-1140 of Kluyveromyces lactis.

GN ORFNames=KLIA0F11143g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E., Barbe V.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe C.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oxtas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RL "Genome evolution in yeasts.";
RN Nature 430:35-44(2004).
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; CR382126; CAG98295.1; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00498; FHA; 2.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00006; FHA DOMAIN; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 813 AA; 91482 MW; 9499356992B8789E CRC64;

Query Match 69.5%; Score 233.5; DB 2; Length 813;
Best Local Similarity 60.3%; Pred. No. 1.9e-21;
Matches 41; Conservative 13; Mismatches 13; Indels 1; Gaps 1;

Qy 1 WFGFRN-PACDYHLGNISRLSNKHFOI-LXXXXXXLLNDISTNGTWTWLNQGVKVNQLL 59
|||||
Db 66 WTFGRNQSCDYHLGDISRLSNKHFOIMLGEGTLMKDKTSTNGTWTWLNQGIQKDTNHILT 125
|||||
Qy 60 QGDEITVG 67
|||||
Db 126 QGDEISVG 133
|||||

RESULT 6
Q6BUB2 PRELIMINARY; PRT; 751 AA.
ID Q6BUB2
AC Q6BUB2;
```



```

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to ep|P22216 Saccharomyces cerevisiae YPL153c SPK1
DE (Fragment).
GN ORFNames=DEHA0C132119;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OC NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Duion B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigie M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boistame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Oztas G.F., Straub M.L., Suleau A.,
RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Sweeney D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/thr protein kinase family.
DR EMBL: CR382135; CAC86283.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR00253; FHA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR00719; Prot_kinase.
DR InterPro: IPR02290; Ser_thr_kinase.
DR InterPro: IPR08271; Ser_thr_pkin_AS.
DR InterPro: IPR08984; SMAD_FHA.
DR InterPro: IPR01245; Tyr_kinase.
DR Pfam: PF00498; FHA; 2.
DR Pfam: PF00089; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00240; FHA; 2.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; Ty_Kc; 1.
DR PROSITE: PS00006; FHA_DOMAIN; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
SQ SEQUENCE 751 AA; 85411 MW; EF8A15F56A788346 CRC64;

Query Match 50.9%; Score 171; DB 2; Length 751;
Best Local Similarity 50.7%; Pred. No. 2.6e-13;
Matches 35; Conservative 9; Mismatches 23; Indels 2; Gaps 1;

Qy 1 WFGFNPACDHYHGNISRLSKHFKQLXXXXXLL--NDISTGTWNGQVKVKNQQL 58
Db 20 WIIRNQECQDYTLNLSTRLSKHFKLWFSNQYTLTQDMSTNGTLNLSRLVKGSNYML 79

Qy 59 SQGDEITVG 67
Db 80 NQGEISVG 88

RESULT 7
Q68329 ID Q68329 PRELIMINARY; PRT; 204 AA.
AC Q68329;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Cds1 Kinase.
GN Name=Cds1;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Ducommun B.;
RT "Evolutionary conservation of a splice variant of the Cds1/Chk2 kinase
RT restricted to its regulatory domain.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ783838; CAH04269.1; -.
DR GO: GO:0016301; F:kinase activity; IEA.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR00719; Prot_kinase.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00240; FHA; 1.
DR PROSITE: PS00006; FHA_DOMAIN; 1.
KW Kinase.
SQ SEQUENCE 204 AA; 23162 MW; 85FC797F553B632D CRC64;

Query Match 40.9%; Score 137.5; DB 2; Length 204;
Best Local Similarity 41.3%; Pred. No. 1.5e-09;
Matches 31; Conservative 13; Mismatches 22; Indels 9; Gaps 2;

Qy 1 WFGFNPACDHYHGNISRLSKHFKQL-----XXXXXXLLNDISTGTWNGQVKV 52
Db 60 WRTGRHKSCVWL-NGPRVSNPFHFIYQGHNRDSENVVFLHDHSSNGTFLNFERLAK 118

Qy 53 NSNQLLSQGDIEITVG 67
Db 119 NSRTILSNGDEIRIG 133

RESULT 8
CDS1_SCHPO ID CDS1_SCHPO STANDARD; PRT; 460 AA.
AC Q09170; O42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase cds1 (EC 2.7.1.37) (Checkpoint kinase
DE Cds1).
GN Name=cds1; ORFNames=SPCC18B5.11c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=95240713; PubMed=7723827; DOI=10.1038/374817a0;
RA Murakami H., Okayama H.;
RT "A kinase from fission yeast responsible for blocking mitosis in S
RT phase.";
RL Nature 374:817-819(1995).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, INTERACTION WITH RAD26, AND
RP PHOSPHORYLATION.

```

RC STRAIN=972;
 RX MEDLINE=98119835; PubMed=9450932;
 RA Lindsey H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
 RA Murray J.M., Osman F., Walworth N., Carr A.M.;
 RT "S-phase-specific activation of Cds1 kinase defines a subpathway of
 RT the checkpoint response in *Schizosaccharomyces pombe*.";
 RL Genes Dev. 12:382-395(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Pearson D., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Fritz C., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Drano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Spakovski G.V., Ussery D., Barrall B.G., Nurse P.;
 RA "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: Has a role in the DNA replication-monitoring S/G2
 CC checkpoint system. It is responsible for blocking mitosis in the S
 CC phase. It monitors DNA synthesis by interacting with DNA
 CC polymerase alpha and sends a signal to block the onset of mitosis
 CC while DNA synthesis is in progress.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
 CC -!- SUBUNIT: Interacts with rad26.
 CC -!- PTM: Autophosphorylated.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. CDS1
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 FHA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X85040; CAA59410.1; -;
 DR EMBL; AJ222869; CAAL1019.1; -;
 DR EMBL; AL109736; CAB52158.1; -;
 DR PIR; S58882; S58882.
 DR HSSP; Q00534; 1JOW.
 DR GeneDB SPombe; SPCC18B5.11c; -;
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008984; SWAD_FHA.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00498; FHA; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00240; FHA; 1.
 DR SMART; SM00240; S_TKC; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Cell cycle; Phosphorylation;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 60 116 FHA.
 FT DOMAIN 167 433 Protein kinase.
 FT NF_BIND 173 181 ATP (By similarity).
 FT BINDING 196 196 ATP (By similarity).
 FT ACT_SITE 294 294 Proton acceptor (By similarity).
 FT CONFLICT 61 61 R -> G (in Ref. 1).
 FT CONFLICT 238 238 F -> C (in Ref. 1).
 SQ SEQUENCE 460 AA; 52014 MW; 4CEB963D3376DB54 CRC64;
 Query Match 40.9%; Score 137.5; DB 1; Length 460;
 Best Local Similarity 41.3%; Pred. No. 3.6e-09;
 Matches 31; Conservative 13; Mismatches 22; Indels 9; Gaps 2;
 Qy 1 WTEGRNPACDYHLGNISRLSNKHQFQL-----XXXXXXLLNDISTNGTWNLGQVK 52
 Db 60 WRFGRHKSCEVWL-NGRVSNFHFYIYQGRNDSSENVFLHDHSSNGTFLNFERLAK 118
 Qy 53 NSNQLLSQGDIEITVG 67
 Db 119 NSRTILSNGDEIRIG 133
 RESULT 9
 Q90ZY5
 ID Q90ZY5 PRELIMINARY; PRT; 503 AA.
 AC Q90ZY5
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Protein Kinase Chk2.
 GN ORFNames-zgc:55865;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chou C.-W., Leu J.-H., Huang C.-J.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AF265346; AAK52419.1; -;
 DR HSSP; P49137; INXK
 DR ZFIN; ZDB-GENE-030131-8942; zgc:55865.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR Pfam; PF00498; FHA; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00240; FHA; 1.
 DR SMART; SM00240; S_TKC; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 503 AA; 56699 MW; BC01BF3DBE4CEC71 CRC64;
 Query Match 34.5%; Score 116; DB 2; Length 503;
 Best Local Similarity 32.4%; Pred. No. 2.5e-06;
 Matches 24; Conservative 11; Mismatches 31; Indels 8; Gaps 1;
 Qy 1 WTEGRNPACDYHLGN-----ISRLSNKHQFQLXXXXXXLLNDISTNGTWNLGQVK 52

```

Db 91 YSFGDRKCDYSFNSILKSPYNTYKSHFRIFRDENLVLDLSNGTWDDEKGN 150
Qy 53 NSNQLSQGBDITV 66
Db 151 GQSLLSNNSVIAL 164

RESULT 10
Q6C079 PRELIMINARY; PRT; 898 AA.
ID Q6C079
AC Q6C079
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to DEHAOC1321lg Debaryomyces hansenii.
GN ORFNames=YALI0F27093g;
OS Yarrowia lipolytica ClrB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talia E.,
RA Goffard N., Prangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckrich J.M., Beyne E., Bleykasten C.,
RA Boistrane A., Boyer E., Fattolico L., Confanieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaki M., Ozias S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.B.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; C382132; CAG78745.1; -.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR008984; SWAD_FHA.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00498; FHA; 2.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 2.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50006; FHA DOMAIN; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 898 AA; 100328 MW; 8C648FB95899D28E CRC64;

```

Query Match

34.2%; Score 115; DB 2; Length 898;

```

Best Local Similarity 34.2%; Pred. No. 6.4e-06;
Matches 25; Conservative 15; Mismatches 27; Indels 6; Gaps 1;
Qy 1 WTFGRNPACDYHLGNISRLSKHFIQI-----LXXXXXXLLNDISTGTWLNQKVKNS 54
Db 82 WTEGRAGSANFVARKARLSKIHFIVEHTGTGKTQKRGVTVIRDSVSLNGTSYNGRYLGRGN 141
Qy 55 NQLLSQGBDITV 67
Db 142 SALLQNGDVITVG 154

RESULT 11
Q803E2 PRELIMINARY; PRT; 503 AA.
ID Q803E2
AC Q803E2
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CHK2 checkpoint homolog.
GN ORFNames=zgc:55865;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Buteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; BC044519; AAH44519.1; -.
DR HSP; P49137; INXK.
DR ZFIN; ZDB-GENE-030131-8942; zgc:55865.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00240; FHA; 1.

```

```

DR SMART, SM00220; S_TKc; 1.
DR PROSITE, PS50006; FHA_DOMAIN; 1.
DR PROSITE, PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 503 AA; 56657 MW; 351A24EA5BBAF8D5 CRC64;

Query Match 33.3%; Score 112; DB 2; Length 503;
Best Local Similarity 31.1%; Pred. No. 8.4e-06;
Matches 23; Conservative 11; Mismatches 32; Indels 8; Gaps 1;

Qy 1 WTGRNPACDYHLGN-----ISRLSNKHFIQLXXXXXXLLNDISTNGTWLNGQKVEK 52
Db 91 YSGFRDRCDYFSNSILKSPYNTYKHKFRIFRDENLVLYEDLSGNGTWVDEKLG 150
Qy 53 NSNQLLSQGDITV 66
Db 151 GKSLGNNNSVIAL 164

RESULT 12
Q87PU0 PRELIMINARY; PRT; 504 AA.
AC Q87PU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein VP1411.
GN OrderedLocusNames=VP1411;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005078; BAC59674.1; -.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 504 AA; 55469 MW; 2F96C75A4F83648E CRC64;

Query Match 31.4%; Score 105.5; DB 2; Length 504;
Best Local Similarity 35.4%; Pred. No. 5.9e-05;
Matches 23; Conservative 14; Mismatches 27; Indels 1; Gaps 1;

Qy 4 GRNPACDYHLGNISR-LSNKHFIQLXXXXXXLLNDISTNGTWLNGQKVEKSNQLLSQGD 62
Db 33 GRAPSCVTLTDHNRFRISGTHCLISVYGDTFYISDVSTNGTMVNGNKLKNQPLISIVEG 92
Qy 63 EITVG 67
Db 93 VVSLG 97

RESULT 13
Q87HC2 PRELIMINARY; PRT; 398 AA.
AC Q87HC2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

```

DE Hypothetical protein VP1043.
GN OrderedLocusNames=VP1043;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005087; BAC62386.1; -.
DR InterPro; IPR000253; FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 398 AA; 44741 MW; 419A55457EB4D173 CRC64;

Query Match 31.1%; Score 104.5; DB 2; Length 398;
Best Local Similarity 33.3%; Pred. No. 6.2e-05;
Matches 23; Conservative 16; Mismatches 27; Indels 3; Gaps 2;

Qy 2 TFCRNPCADYHLGNISR-LSNKHFIQLXXXXXXLLNDISTNGTWLNG--QKVEKSNQLL 58
Db 33 TFCRSEQCDWTLDPDPERVISGTHGELIKFGDKYLKDLSTNGTFVNNAVPIQGNEAL 92
Qy 59 SQGDEITVG 67
Db 93 SHGDTVALG 101

RESULT 14
Q98IK6 PRELIMINARY; PRT; 359 AA.
AC Q98IK6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mr2358 protein.
GN OrderedLocusNames=mlr2358;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303059;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002399; BAB49510.1; -.
DR InterPro; IPR000253; FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 359 AA; 39374 MW; 36732CCBB6AB7DD5 CRC64;

Query Match 29.9%; Score 100.5; DB 2; Length 359;

```

Matches	26;	Conservative	4;	Mismatches	36;	Indels	2;	Gaps	1;
Qy	1	WTGRNPACDYHLGNISRLSNKHFIQL--XXXXXXLLNDISTNGTWNLGOKVEKNSQLL	58						
Dd	34	WTIGRKKACDPSFGNKLVSGEHCKITVNEESGSVLEDTSTNGTVINKLKVVKKQTPL	93						
Qy	59	SQGDEITY 66							
Dd	94	KNGDVIVY 101							
 RESULT 16									
Q89P85	ID	Q89P85	PRELIMINARY;	PRT;	460 AA.				
AC	Q89P85;								
DT	01-JUN-2003 (TEMBLrel. 24, Created)								
DT	01-JUN-2003 (TEMBLrel. 24, Last sequence update)								
DT	01-MAR-2004 (TEMBLrel. 26, Last annotation update)								
DE	Bir3598 protein.								
GN	OrderedLocusNames=blr3598;								
OS	Bradyrhizobium japonicum.								
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;								
OC	Bradyrhizobiaceae; Bradyrhizobium.								
OX	NCBI_TaxID=375;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=USDA110;								
RX	MEDLINE=22484998; PubMed=12597275;								
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchimi T.,								
RA	Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,								
RA	Kohara M., Matsumoto M., Shimpō S., Teiruoka H., Wada T., Yamada M.,								
RA	Tabata S.;								
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium								
RT	Bradyrhizobium japonicum USDA110.";								
RL	DNA Res. 9:189-197(2002).								
DR	EMBL: AP005948; BAC48863.1; -.								
DR	InterPro: IPR00253; FHA.								
DR	InterPro: IPR008984; SMAD FHA.								
DR	Pfam: PF00498; FHA; 1.								
DR	SMART: SM00240; FHA; 1.								
DR	PROSITE: PS50006; FHA_DOMAIN; 1.								
KW	Complete proteome.								
SQ	SEQUENCE 460 AA; 50895 MW; C2CC7E01310D367F CRC64;								
 Query Match 29.3%; Score 98.5; DB 2; Length 460;									
Best Local Similarity 35.4%; Pred.No. 0.00044;									
Matches	23;	Conservative	12;	Mismatches	29;	Indels	1;	Gaps	1;
Qy	4	GRNPACDYHLGNISR-LSNKHFIQLXXXXXXLLNDISTNGTWNLGOKVEKNSQLLSQGD	62						
Dd	33	GRDLHLDTLTPDPARTISGHCVEHFRRDGGYWLHDYSTNGTFLNGADQRMGRPRLRDGD	92						
Qy	63	EITVG 67							
Dd	93	RLTIG 97							
 RESULT 17									
Q8DJ88	ID	Q8DJ88	PRELIMINARY;	PRT;	1029 AA.				
AC	Q8DJ88;								
DT	01-MAR-2003 (TEMBLrel. 23, Created)								
DT	01-MAR-2003 (TEMBLrel. 23, Last sequence update)								
DT	01-MAR-2004 (TEMBLrel. 26, Last annotation update)								
DE	ABC transporter ATP-binding protein.								
GN	OrderedLocusNames=cil1340;								
OS	Synechococcus elongatus (Thermosynechococcus elongatus).								
OC	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.								
OX	NCBI_TaxID=32046;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=BP-1;								
RX	MEDLINE=22225144; PubMed=12240834;								

RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.,
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1",
RT DNA Res. 9:123-130(2002).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AP005373; BAC08892.1; -.
DR HSPSP; P46013; IR21.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_Transporter.
DR InterPro; IPR000253; FHA_Transporter.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00498; FHA; 2.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00240; FHA; 3.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 1029 AA; 114450 MW; E4B6278B017BD179 CRC64;

Query Match 28.3%; Score 95; DB 2; Length 1029;
Best Local Similarity 38.2%; Pred. No. 0.003;
Matches 26; Conservative 15; Mismatches 23; Indels 4; Gaps 3;

Qy 1 WTGRNPACDYHLGNISRLSKVHFOILXXXXXLLNDI-STNGTWLNGQVKVKNQSLLS 59
Db 260 FTIGRDPNNDLVIGH-PTVSRHHAKIERRNGDFLLTDLGSSNGTFVNGREVEEPT--LLR 316

Qy 60 QGDEITVG 67
Db 317 VGDSIRIG 324

RESULT 18
Q9FWAO PRELIMINARY; PRT; 567 AA.
ID Q9FWAO
AC Q9FWAO; TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F11A12.9 (Fragment).
GN Name=F11A12.9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Erosoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Rinning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.

```
Db 19 SLGFKPGSTIRIGRIVRGNEIAIKDAGISTKHLRIVSDSENWIHDLGSSNGTILNSDTI 78
Qy 51 EKSNQILLSQDEITVG 67
Db 79 DSOTFVNLSHGDEIKLG 95

RESULT 20
Q6BH59 PRELIMINARY; PRT; 443 AA.
ID Q6BH59
AC Q6BH59;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to CA4770|PF9382.3 Candida albicans IPF9382.3.
GN ORFNames=DEHROG22409;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Lafontaine I., de Montigny J., Warck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer E., Faiched C., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Faiched C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; CR382139; CAG90972.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008277; Ser_thr_kin_AS.
DR InterPro; IPR008984; SWAD_FHA.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00498; FHA; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 443 AA; 50501 MW; 80B5BACD8B8AB9E9D CRC64;

Query Match 27.5%; Score 92.5; DB 2; Length 443;
Best Local Similarity 32.3%; Pred. No. 0.0026;

Matches 26; Conservative 13; Mismatches 25; Indels 10; Gaps 3;

Qy 3 FGRNPACDYHLGNISR-LSNKHFOILXXXXXXLLNDISTNGTWL-----NQCKVEKN 53
Db 3 FGRSESCDWLPDPERIISGVHGEITKFGNDYLLDLSTNGIFVYKNSVSPVGVNGVEVALN 62

Qy 54 SNOLLSQGD-EITV 66
Db 63 DKDVINFGDYIEIV 76

RESULT 22
Q7MDP1 PRELIMINARY; PRT; 434 AA.
ID Q7MDP1
AC Q7MDP1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Uncharacterized conserved protein.
GN OrderedLocusNames=VVA0995;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
```

```
Matches 21; Conservative 18; Mismatches 21; Indels 5; Gaps 2;

Qy 4 GRNPACDYHLGNISRLSNKHFOILXX-----XXXLLNDISTNGTWLNGQVKNSQLLS 59
Db 13 GRSSSCDVRI-NGNDVSSRKCKLTLTINNREYLCIKDLSNGTYLNDIIGKSSILLR 71

Qy 60 QGDEI 64
Db 72 SGDKL 76

RESULT 21
Q8D6T0 PRELIMINARY; PRT; 374 AA.
ID Q8D6T0
AC Q8D6T0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Uncharacterized conserved protein.
GN OrderedLocusNames=VV20445;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE016809; AAC07399.1; -.
DR InterPro; IPR00253; FHA_FHA.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 42554 MW; F42608D67A766FE9 CRC64;

Query Match 27.1%; Score 91; DB 2; Length 374;
Best Local Similarity 35.1%; Pred. No. 0.0034;

Matches 26; Conservative 13; Mismatches 25; Indels 10; Gaps 3;

Qy 3 FGRNPACDYHLGNISR-LSNKHFOILXXXXXXLLNDISTNGTWL-----NQCKVEKN 53
Db 3 FGRSESCDWLPDPERIISGVHGEITKFGNDYLLDLSTNGIFVYKNSVSPVGVNGVEVALN 62

Qy 54 SNOLLSQGD-EITV 66
Db 63 DKDVINFGDYIEIV 76

RESULT 22
Q7MDP1 PRELIMINARY; PRT; 434 AA.
ID Q7MDP1
AC Q7MDP1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Uncharacterized conserved protein.
GN OrderedLocusNames=VVA0995;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
```

RESULT 25


```
DR SMART; SW00240; FHA; 1.
KW PROSITE; PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 159 AA; 17667 MW; F4F44A3D3BF05AA1 CRC64;

Query Match          25.9%; Score 87; DB 2; Length 159;
Best Local Similarity 37.3%; Pred. No. 0.0044;
Matches 25; Conservative 9; Mismatches 29; Indels 4; Gaps 3;

QY      2 TFGRNPAACYHLGNISRLSNKHPQLIXLXXXXXLLNDI-STNGTWLNQKVKNLSQ 60
       |||||:::|||||:::|||||:::|||||:::|||||:::
Db      88 TIGCKDNSTML-NEGVSGHARVYLNNQYILEDINSLNQTGTVLNQKIK--SKAYIKS 144

QY      61 GDEITVG 67
       |||||:::
Db      145 GDBIKIG 151

RESULT 27
MEKI SCHPO
ID MEKI SCHPO STANDARD; PRT; 445 AA.
AC Q10292;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Meiosis-specific serine/threonine-protein kinase mek1 (BC 2.7.1.37).
GN Namesmek1; ORFNames=SPAC14C4.03;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
[1]
RP SEQUENCE FROM N.A.
RA Lyne M.H., Bryant J.A., Aves S.J.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA STRAIN=972;
MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Scourts J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,
RA James K., Jones L., Jones M., Leather S., McDonald D., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabriel C., Fuhs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe."
RL Nature 415:871-880(2002).
CC -! FUNCTION: Probable protein kinase required for meiotic recombination.
CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -! SIMILARITY: Contains 1 FHA domain.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; Z71478; CAB96101.1; --
DR EMBL; Z98596; CAB1196.1; --
DR PIR; T43420; T43420.
DR HSSP; P49137; IKWP
DR GeneDB_Sfombe; SPAC14C4.03; --
DR InterPro; IPR000253; FHA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOW; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Meiosis; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 62 116 FHA.
FT DOMAIN 160 421 Protein kinase.
FT NP_BIND 166 174 ATP (By similarity).
FT ACT_SITE 281 281 proton acceptor (By similarity).
FT BINDING 189 189 ATP (By similarity).
SQ SEQUENCE 445 AA; 51151 MW; 6062D79E9A1B45B0 CRC64;

Query Match 25.7%; Score 86.5; DB 1; Length 445;
Best Local Similarity 33.3%; Pred. No. 0.016;
Matches 23; Conservative 14; Mismatches 25; Indels 7; Gaps 2;

QY 4 GRNPACDYHLGNISRLSNKHQFQILXXXXXXLLN-----DISTNGTWLNGQKVEKNSQL 57
DB 65 GRNTCNQYLQFT-ASYKHFRVYLVLDMDPLVYCDQSSNGTFLNRLRGKNSVL 123
QY 58 LSGDEITV 66
DB 124 LSGDILDV 132

RESULT 28
Q91751 PRELIMINARY; PRT; 497 AA.
AC Q91751
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA0081;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Muzoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004447; AAG03471.1; --
DR PIR; F83634; F83634.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SWAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 497 AA; 53310 MW; A6C63E1D287E7272 CRC64;
Query Match 25.4%; Score 85.5; DB 2; Length 497;
Best Local Similarity 33.3%; Pred. No. 0.024;
Matches 23; Conservative 13; Mismatches 30; Indels 3; Gaps 2;
QY 2 TGRNPACDYHLGNISRL-SNKHQFQILXXXXXXLLNDISNGTWL--NGQKVEKNSQL 58
DB 29 TIGRGPNDWLPDPRLVSSRHCTILNRDGVVYLTDTSTNGVLLVNAHRLRGNSEPL 88
QY 59 SGDEITV 67
DB 89 QDGETVRLG 97
RESULT 29
CHFR_HUMAN STANDARD; PRT; 664 AA.
ID_CHFR_HUMAN Q96FL1; Q96SL3; Q9NRT4; Q9NT32; Q9NVD5;
AC Q96FL1; Q96SL3; Q9NRT4; Q9NT32; Q9NVD5;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ubiquitin ligase protein CHFR (EC 6.3.2.-) (Checkpoint with forkhead
DE and RING finger domains protein).
GN Name=CHFR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, AND
RP VARIANT MET-580.
RX MEDLINE=203988685; PubMed=10935642; DOI=10.1038/35019108;
RA Scolnick D.M., Halazonetis T.D.;
RT "Chfr defines a mitotic stress checkpoint that delays entry into
RT metaphase".
RL Nature 406:430-435(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Teratocarcinoma;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
Omura Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyama S., Komai P., Harai R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama K., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT VAL-497.
RP TTSUE=Placenta;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derce J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Stenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 359-664 FROM N.A.
RP TTSUE=Testis;
RC Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN FUNCTION, AUTOUBIQUITINATION, AND MUTAGENESIS OF ILE-306 AND TRP-332.
RX PubMed=11807090; DOI=10.1083/jcb.200108016;
RA Kang D., Chen J., Wong J., Fang G.;
RT "The checkpoint protein Chfr is a ligase that ubiquitinates Plk1 and
RT inhibits Cdc2 at the G2 to M transition.";
RL J. Cell Biol. 156:249-259(2002).
RN [6]
RN FUNCTION, AUTOUBIQUITINATION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL
RP STAGE.
RX PubMed=11912157;
RA Chaturvedi P., Sudakin V., Bobiak M.L., Fisher P.W., Mattern M.R.,
RA Jablonski S.A., Hurle M.R., Zhu Y., Yen T.J., Zhou B.-B.;
RT "Chfr regulates a mitotic stress pathway through its RING-finger
RT domain with ubiquitin ligase activity.";
RL Cancer Res. 62:1797-1801(2002).
RN [7]
RN DISEASE, AND VARIANTS ARG-270; VAL-497 AND MET-580.
RX PubMed=11948416; DOI=10.1038/sj/ncr.1205402;
RA Mizuno K., Osada H., Konishi H., Tatematsu Y., Yatabe Y.,
RA Mitsudomi T., Fujii Y., Takahashi T.;
RT "Aberrant hypermethylation of the Chfr prophase checkpoint gene in
RT human lung cancers.";
RL Oncogene 21:2328-2333(2002).
RN [8]
RN DISEASE.
RX PubMed=12538348; DOI=10.1093/carcin/24.1.47;
RA Corn P.G., Summers M.K., Fogt F., Virmani A.K., Gazdar A.F.,
RA Halazonetis T.D., El-Deiry W.S.;
RT "Frequent hypermethylation of the 5' CpG island of the mitotic stress
RT checkpoint gene Chfr in colorectal and non-small cell lung cancer.";
RL Carcinogenesis 24:47-51(2003).
RN [9]
RN DISEASE.

RX PubMed=12810945; DOI=10.1073/pnas.13370661100;
RA Toyota M., Sasaki Y., Satoh A., Ogi K., Kikuchi T., Suzuki H.,
RA Mita H., Tanaka N., Itoh F., Issa J.-P.J., Jair K.-W., Schuebel K.E.,
RA Imai K., Tokino T.;
RT "Epigenetic inactivation of Chfr in human tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7818-7823(2003).
RN [10]
RN DISEASE.
RX PubMed=14695171;
RA Satoh A., Toyota M., Itoh F., Sasaki Y., Suzuki H., Ogi K.,
RA Kikuchi T., Mita H., Yamashita T., Kojima T., Kusano M., Fujita M.,
RA Hosokawa M., Endo T., Tokino T., Imai K.;
RT "Epigenetic inactivation of Chfr and sensitivity to microtubule
RT inhibitors in gastric cancer.";
RL Cancer Res. 63:8606-8613(2003).
RN [11]
RN DISEASE, AND VARIANTS NSCLC LEU-166; PRO-202 AND SER-536.
RX PubMed=14612512;
RA Mariatos G., Bothos J., Zacharatos P., Summers M.K., Scolnick D.M.,
RA Kittas C., Halazonetis T.D., Gorgoulis V.G.;
RT "Inactivating mutations targeting the chfr mitotic checkpoint gene in
RT human lung cancer.";
RL Cancer Res. 63:7185-7189(2003).
RN [12]
RN PHOSPHORYLATION, AND MUTAGENESIS OF THR-39 AND SER-205.
RX PubMed=14638868;
RA Schivelman E.;
RT "Promotion of mitosis by activated protein kinase B after DNA damage
RT involves polo-like kinase 1 and checkpoint protein Chfr.";
RL Mol. Cancer Res. 1:959-969(2003).
RN [13]
RN FUNCTION, INTERACTION WITH UBE2V2, AND PHOSPHORYLATION.
RX PubMed=14562038; DOI=10.1038/sj.onc.1206831;
RA Bothos J., Summers M.K., Venero M., Scolnick D.M., Halazonetis T.D.;
RT "The Chfr mitotic checkpoint protein functions with Ubc13-Mms2 to form
RT Lys63-linked polyubiquitin chains.";
RL Oncogene 22:7101-7107(2003).
RN [14]
RN FUNCTION.
RX PubMed=14694445; DOI=10.1002/mc.10161;
RA Erson A.B., Petty E.M.;
RT "Chfr-associated early G2/M checkpoint defects in breast cancer
RT cells.";
RL Mol. Carcinog. 39:26-33(2004).
RN [15]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 14-128.
RX PubMed=12121644; DOI=10.1016/S0969-2126(02)00776-1;
RA Stavridi E.S., Huyen Y., Loreto I.R., Scolnick D.M., Halazonetis T.D.,
RA Pavletich N.P., Jeffrey P.D.;
RT "Crystal structure of the FHA domain of the Chfr mitotic checkpoint
RT protein and its complex with tungstate.";
RL Structure 10:891-899(2002).
RN CC -1- FUNCTION: E3 ubiquitin-protein ligase required to transiently
CC arrest cells in early prophase when they are exposed to
CC microtubule poisons. Acts in early prophase before chromosome
CC condense, when the centrosome move apart from each other along the
CC periphery of the nucleus. Probably promotes the formation of Lys-
CC 63-linked polyubiquitin chains and functions with the specific
CC ubiquitin-conjugating UBC13-MMS2 (UBE2N-UBE2V2) heterodimer.
CC Substrates that are polyubiquitinated at Lys-63 are usually not
CC targeted for degradation, but are rather involved in signaling
CC cellular stress. This suggests that it may be involved in
CC signaling the presence of mitotic stress caused by microtubule
CC poisons.
RN CC -1- PATHWAY: Ubiquitin conjugation; third step.
RN CC -1- SUBCELLULAR LOCATION: Nuclear.
RN CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q96EP1-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96EP1-2; Sequence=VSP_009349;
CC Name=3;

```

CC IsoId=Q96P1-3; Sequence=VSP_009350;
CC Note=No experimental confirmation available;
CC -I- TISSUE SPECIFICITY: Ubiquitous.
CC -I- DEVELOPMENTAL STAGE: Weakly expressed in G1 phase, and highly
CC expressed during S phase.
CC -I- DOMAIN: The FHA domain may be required to interact with
CC phosphorylated proteins.
CC -I- PFM: Autoubiquitinated in vitro.

Query Match          25.3%; Score 85; DB 1; Length 664;
Best Local Similarity 34.8%; Pred. No. 0.038;
Matches 23; Conservative 6; Mismatches 35; Indels 2; Gaps 1;

Qy 1 WFGNRNACDYLHGNISRLSNKHFIQIL--XXXXXXLLNDISTNGTWLNGQVKENSNQLL 58
    ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 38 WTIGRRGCDLSPFSNKLIVGDHCRIVVDKSGQVTLDTSTSGTVINKLKVKKQTCPL 97
    ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 59 SQGDEI 64
    ||| |||
Db 98 QTGDVI 103
    ||| |||

RESULT 30
Q6MQJ2 PRELIMINARY; PRT; 674 AA.
ID Q6MQJ2 AC Q6MQJ2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Bd0474;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842647; CAE78455.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00498; FHA; 1.
DR SMART; SMO0240; FHA; 2.
DR PROSITE; PS50006; FHA_DOMAIN; 2.
DR Complete proteome.
KW SEQUENCE 674 AA; 75694 MW; E898524PE1A87766 CRC64;

Query Match          25.3%; Score 85; DB 2; Length 674;
Best Local Similarity 29.9%; Pred. No. 0.039;
Matches 20; Conservative 12; Mismatches 33; Indels 2; Gaps 2;

Qy 1 WFGNRNACDYLHGNISRLSNKHFIQILXXXXXXLLNDI-STNGTWLNGQVKENSNQLS 59
    ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 165 WIAGRSTCHIQIRD-QRVSRQFEIRAGSQFVILDLGVSNGTLNGNPISSTDWTFIK 223
    ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 60 QGDEIV 66
    ||| |||
Db 224 SGDAIV 230
    ||| |||

RESULT 31
Q7UIF2 PRELIMINARY; PRT; 346 AA.
ID Q7UIF2 AC Q7UIF2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

```

CC		the degradation of substrates such as PLK, a protein kinase involved in mitotic progression. However, as experiments have been done either in vitro or with extracts from Xenopus, there is actually few evidence for a role of CHFR in protein degradation in vivo.
CC		-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC		-----
CC	EMBL; AK052473; BAC35008.1; -	
CC	EMBL; AK077629; BAC36912.1; -	
DR	EMBL; BC049792; AAH49792.1; -	
DR	HSSP; Q96EP1; ILGP	
DR	MCD; MGI:2444898; 573048M20Rik.	
DR	InterPro; IPR000253; FHA.	
DR	InterPro; IPR008984; SMAD_FHA.	
DR	InterPro; IPR001841; Znf_Ring.	
DR	Pfam; PF004498; FHA; 1.	
DR	Pfam; PR000397; zf-C3HC4; 1.	
DR	SMART; SMO0240; FHA; 1.	
DR	SMART; SMO0184; RING; 2.	
DR	PROSITE; PS50006; FHA_DOMAIN; 1.	
DR	PROSITE; PS00518; ZF_RING_1; 1.	
DR	PROSITE; PS50089; ZF_RING_2; 1.	
KW	Alternative splicing; Cell cycle; Ligase; Metal-binding; Mitosis;	
KW	Nuclear protein; Phosphorylation; UbI conjugation;	
KW	UbI conjugation pathway; Zinc; Zinc-finger.	
FT	DOMAIN 38 89	FHA.
FT	ZN_FING 303 342	RING-type.
FT	VARSPLIC 470 470	Missing (in isoform 2).
FT		/FTId=VSP_009351.
FT	CONFLICT 98 98	Q -> H (in Ref. 1).
FT	CONFLICT 531 531	G -> C (in Ref. 1).
FT	SEQUENCE 664 AA; 73871 MW; D6S1BE3E463DEBB6 CRC64;	
SQ		
	Query Match 25.0%; Score 84; DB 1; Length 664;	
	Best Local Similarity 33.3%; Pred. No. 0.051;	
	Matches 22; Conservative 6; Mismatches 36; Indels 2; Gaps 1;	
QY	1 WTGRNPACDYHLGNTSRLSRKHFQL--XXXXXXLLINDISTNGTWLNGOKVERNSQLL 58 : : : : : : : : : : : : : : : : : :	
Dd	38 WTIERRRGCDLSPFNKLVSGDCKLTVDISEGVLEDTSTGTINKLVQVKQKYPL 97 : : : : : : : : : : : : : : : : : :	
QY	59 SQGDEI 64	
Dd	98 QSGDII 103	
RESULT 33		
QNHD4	ID	PRELIMINARY; PRT; 143 AA.
ID	QNHD4	
AC	QNHD4;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	03-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	03-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Putative signal transduction protein.	
GN	OrderedLocusNames=DIP1204;	
OS	Corynebacterium diphtheriae.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OX	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.	
NCBI_TaxID=171;	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAN=Blototype gravis / NCTC 13129:	
RX	MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;	
RA	Cardeno-Parraga A.-M., Estratouliou A., Dover L.G., Holden M.T.G.,	
RA	Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,	
RA	De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,	

Db 272 VITVG 276

RESULT 35

Q8FUH9

ID Q8FUH9 PRELIMINARY; PRT; 298 AA.

AC Q8FUH9; (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=CE0039;

OS Corynebacterium efficiens.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.

OX NCBI_TaxID=152794;

RN [1]_TaxID=152794;

RP SEQUENCE FROM N.A.

RC STRAIN=YS-314;

RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;

RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,

RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Iike K.,

RA Gojobori T.;

RT "Comparative complete genome sequence analysis of the amino acid

RT replacements responsible for the thermostability of Corynebacterium

RT efficiens.";

RL Genome Res. 13:1572-1579(2003).

DR EMBL; AP05214; BAC16849.1; --

DR HSSP; P46014; 1M2K.

DR InterPro; IPR00253; FHA.

DR InterPro; IPR008984; SMAD_FHA.

DR Pfam; PF00498; FHA; 1.

DR SMART; SM00240; FHA; 1.

DR PROSITE; PS50006; FHA_DOMAIN; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 298 AA; 32166 MW; ACAD78F03746781F CRC64;

Query Match 24.9%; Score 83.5; DB 2; Length 298;

Best Local Similarity 40.0%; Pred. No. 0.025;

Matches 26; Conservative 10; Mismatches 24; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNOLLSSGD 62

Db 227 GRSNDADLRPD-TGVSQRHAEITWDGRDAILVDLKTNGTTPVE---NWLLADGD 282

QY 63 EITVG 67

Db 283 VITVG 287

RESULT 36

Q6MQJ4

ID Q6MQJ4 PRELIMINARY; PRT; 557 AA.

AC Q6MQJ4;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=Bd0472;

OS Bdellovibrio bacteriovorus.

OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;

OC Bdellovibrionaceae; Bdellovibrio.

OX NCBI_TaxID=959;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;

RX PubMed=14752164; DOI=10.1126/science.1093027;

RA Rudulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,

RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,

RA Sockett R.E., Schuster S.C.;

RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a

RT genomic perspective.";

RL Science 303:689-692(2004).

Query Match 24.9%; Score 83.5; DB 2; Length 288;

Best Local Similarity 38.5%; Pred. No. 0.024;

Matches 25; Conservative 10; Mismatches 25; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNOLLSSGD 62

Db 216 GRNEAHLRPD-TGVSQRHAEITWNGRDAILTDLKTNGTTPVE---NWLLDGD 271

QY 63 EITVG 67

Db 125 GDEVOIG 131

RESULT 34

Q6NKG7

ID Q6NKG7 PRELIMINARY; PRT; 288 AA.

AC Q6NKG7;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=DIP0059;

OS Corynebacterium diptheriae.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.

OX NCBI_TaxID=1711;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Biotype Gravis / NCTC 13129;

RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;

RA Cardeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,

RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,

De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,

RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,

RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,

RA Whitehead S., Barrell B.G., Parkhill J.;

RT "The complete genome sequence and analysis of Corynebacterium

RT diptheriae NCTC13129.";

RL Nucleic Acids Res. 31:6516-6523(2003).

DR EMBL; BX248354; CAE48567.1; --

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SMAD_FHA.

DR Pfam; PF00498; FHA; 1.

DR SMART; SM00240; FHA; 1.

DR PROSITE; PS50006; FHA_DOMAIN; 1.

KW Complete proteome.

SQ SEQUENCE 288 AA; 31558 MW; AB8B52C505D89241 CRC64;

Query Match 24.9%; Score 83.5; DB 2; Length 288;

Best Local Similarity 38.5%; Pred. No. 0.024;

Matches 25; Conservative 10; Mismatches 25; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNOLLSSGD 62

Db 216 GRNEAHLRPD-TGVSQRHAEITWNGRDAILTDLKTNGTTPVE---NWLLDGD 271

QY 63 EITVG 67

Db 125 GDEVOIG 131

RESULT 34

Q6NKG7

ID Q6NKG7 PRELIMINARY; PRT; 288 AA.

AC Q6NKG7;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=DIP0059;

OS Corynebacterium diptheriae.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.

OX NCBI_TaxID=1711;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Biotype Gravis / NCTC 13129;

RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;

RA Cardeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,

RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,

De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,

RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,

RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,

RA Whitehead S., Barrell B.G., Parkhill J.;

RT "The complete genome sequence and analysis of Corynebacterium

RT diptheriae NCTC13129.";

RL Nucleic Acids Res. 31:6516-6523(2003).

DR EMBL; BX248354; CAE48567.1; --

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SMAD_FHA.

DR Pfam; PF00498; FHA; 1.

DR SMART; SM00240; FHA; 1.

DR PROSITE; PS50006; FHA_DOMAIN; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 288 AA; 31558 MW; AB8B52C505D89241 CRC64;

Query Match 24.9%; Score 83.5; DB 2; Length 288;

Best Local Similarity 38.5%; Pred. No. 0.024;

Matches 25; Conservative 10; Mismatches 25; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSKHFIQLXXXXXXLLNDI-STNGTWLNGQKVEKNSNOLLSSGD 62

Db 216 GRNEAHLRPD-TGVSQRHAEITWNGRDAILTDLKTNGTTPVE---NWLLDGD 271

QY 63 EITVG 67

Db 125 GDEVOIG 131

RESULT 34

Q6NKG7

ID Q6NKG7 PRELIMINARY; PRT; 288 AA.

AC Q6NKG7;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=DIP0059;

OS Corynebacterium diptheriae.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.

OX NCBI_TaxID=1711;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Biotype Gravis / NCTC 13129;

RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;

RA Cardeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,

RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,

De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,

RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A

```
DR EMBL: BX842647: CAE78453.1; -.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR SMART: SM00240; FHA; 1.
DR PROSITE: PS50006; FHA_DOMAIN; 1.
KW Complete proteome.
SQ SEQUENCE 557 AA; 60066 MW; C057EB85BP578A20 CRC64;

Query Match      24.6%; Score 82.5; DB 2; Length 557;
Best Local Similarity 34.4%; Pred. No. 0.067;
Matches 22; Conservative 15; Mismatches 22; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSNKHFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQGD 62
Db 26 GRAPQCCDIKLIS-SGVSKHEITEIAVFKDKIIVTDLRSSNGTYLNGVRVQSGVMRL---GD 81

QY 63 EITV 66
Db 82 KLGI 85

RESULT 37
Q8FTJ5 PRELIMINARY; PRT; 142 AA.
AC Q8FTJ5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CE1573;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y8-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL: AP005219; BAC18383.1; -.
DR HSP: P46014; IMZK.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR SMART: SM00240; FHA; 1.
DR PROSITE: PS50006; FHA_DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 142 AA; 15096 MW; 97C64454E33A1238 CRC64;

Query Match      24.3%; Score 81.5; DB 2; Length 142;
Best Local Similarity 32.8%; Pred. No. 0.021;
Matches 22; Conservative 15; Mismatches 25; Indels 5; Gaps 3;

QY 2 TFGNPACDYHLGNISRLSNKHFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQ 60
Db 68 TAGRHPSDFLDVTVT-VSRHAEFRINEGEFEYVDVGSGLNGTYVNR---EPRNSQVLQT 123

QY 61 GDEITVG 67
Db 124 GDEIQIG 130

RESULT 38
Q8NU92 PRELIMINARY; PRT; 288 AA.
ID Q8NU92
AC Q8NU92
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Cgl0047.
GN OrderedLocusNames=Cgl0047;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005274; BAB97440.1; -.
DR HSP: P46014; IMZK.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR SMART: SM00240; FHA; 1.
DR PROSITE: PS50006; FHA_DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 288 AA; 31317 MW; 1E0B67A6F95AA791 CRC64;

Query Match      24.3%; Score 81.5; DB 2; Length 288;
Best Local Similarity 38.5%; Pred. No. 0.044;
Matches 25; Conservative 11; Mismatches 24; Indels 5; Gaps 3;

QY 4 GRNPACDYHLGNISRLSNKHFQILXXXXXXLLNDI-STNGTWNLGQKVEKNSNQLLSQGD 62
Db 217 GRSDADLRLPD-TGVSQRHQHVEITWGRDAILDLKSTNGTVDTPVD---NWLADGD 272

QY 63 EITVG 67
Db 273 VITVG 277

RESULT 39
Q6M8V2 PRELIMINARY; PRT; 298 AA.
ID Q6M8V2
AC Q6M8V2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Cg0064;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.P., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
DR EMBL: BX927148; CAF18615.1; -.
DR InterPro: IPR000253; FHA.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR SMART: SM00240; FHA; 1.
DR PROSITE: PS50006; FHA_DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 298 AA; 32347 MW; 0BF6A50BA2448EF2 CRC64;
```

```

Query Match          24.3%; Score 81.5; DB 2; Length 298;
Best Local Similarity 38.5%; Pred.No.0.046;
Matches 25; Conservative 11; Mismatches 24; Indels 5; Gaps 3;

QY      4 GRNPACDYHLGNISRLSNKHFQILXXXXXXXXLLNDI-STNGTWNLGQVKVKNSQLLSOGD 62
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      227 GRNDDADLRUPD-TGVSQRQHVEITWGRDAILLDLKSTNGTNTVDTPD---NWLLADGD 282
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY      63 EITVG 67
         |||
Db      283 VIIVG 287
         |||

RESULT 40
YKI5 CAEEL
ID_YKI5 CAEEL STANDARD; PRT; 952 AA.
AC P46012;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein C01G6.5 in chromosome III.
GN CRFNames=C01G6.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
   RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -1- SIMILARITY: Contains 1 FHA domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
-----
DR ENBL; Z35595; CAA84636.1; -.
DR PIR; T18837; T18837.
DR WormBase; WBGene00007227; C01G6.5.
DR WormPeP; C01G6.5; CE00867.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00499; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PSS0006; FHA_DOMAIN; 1.
DR KW Hypothetical protein.
FT DOMAIN 39 100 FHA
SQ SEQUENCE 952 AA; 105274 MW; D9CFB0AB3C685FBE CRC64;

Query Match          24.1%; Score 81; DB 1; Length 952;
Best Local Similarity 41.2%; Pred.No.0.19;
Matches 14; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY      34 LLNDISTNGTWNLGQVKVKNSQLLSOGDEITVG 67
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      84 LVEDISNGYIINDRRLSKDKREILKSGDIKFG 117

```